PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, 15/16, 15/18, 15/63, C07K 14/46, 14/47, 14/475

A1

(11) International Publication Number: WO 99/07850

(43) International Publication Date: 18 February 1999 (18.02.99)

(21) International Application Number: PCT/US98/16502

(22) International Filing Date: 6 August 1998 (06.08.98)

(30) Priority Data:
60/054,966 6 August 1997 (06.08.97) US
60/058,108 5 September 1997 (05.09.97) US

(71) Applicant: MILLENNIUM BIOTHERAPEUTICS, INC. [US/US]; 620 Memorial Drive, Cambridge, MA 02142 (US).

(72) Inventors: HOLTZMAN, Douglas, A.; 821 Centre Street #6, Jamaica Plain, MA 02130 (US). GOODEARL, Andrew, D., J.; 13 Indian Ridge Road, Natick, MA 01760 (US).

(74) Agent: MEIKLEJOHN, Anita, L.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83 NUCLEIC ACID MOLECULES AND POLYPEPTIDES

(57) Abstract

The invention relates to Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 polypeptides, nucleic acid molecules encoding Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83, and uses thereof.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83 NUCLEIC ACID MOLECULES AND POLYPEPTIDES

Cross Reference to Related Applications

This application claims priority from U.S. Serial Number 60/054,966, filed August 6, 1997 and U.S. Serial Number 60/058,108, filed September 5, 1997.

Summary of the Invention

The invention relates to the discovery and characterization of Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83. Tango-71 is a human protein which is approximately 90% identical to murine ADAMTS-1. Tango-73 is a human protein that is 48% identical to rate RVP.1 (Briehl et al., *Mol. Endocrinol.* 5:1381, 1991). Tango-74 is a human protein with homology to TRAIL receptor (Pan et al., *Science* 276:111, 1997). Tango-76 is a rat protein which is approximately 40% identical to murine ADAMTS-1. Tango-83 is expressed by stimulated human astrocytes.

The invention features isolated nucleic acid molecules encoding Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptides; isolated nucleic acid molecules encoding polypeptides which are substantially similar to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83; and isolated nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule having the sequence of the protein coding portion of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:9.

The invention also features a host cell which includes an isolated nucleic acid molecule encoding Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 and a nucleic acid vector (e.g., an expression vector; a vector which includes a regulatory element; a vector which includes a regulatory element selected from the group consisting of the cytomegalovirus hCMV immediate early

gene, the early promoter of SV40 adenovirus, the late promoter of SV40 adenovirus, the lac system, the trp system, the TAC system, the TRC system, the major operator and promoter regions of phage λ, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase, the promoters of acid phosphatase, and the promoters of the yeast α-mating factors; vector which includes a regulatory element which directs tissue-specific expression; a vector which includes a reporter gene; a vector which includes a reporter gene selected from the group selected from the group consisting of β-lactamase, chloramphenicol acetyltransferase (CAT), adenosine deaminase (ADA), aminoglycoside phosphotransferase (neo^r, $G418^r$), dihydrofolate reductase (DHFR), hygromycin-B-phosphotransferase (HPH), thymidine kinase (TK), lacZ (encoding β-galactosidase), and xanthine guanine phosphoribosyltransferase (XGPRT); a vector that is a plasmid, a vector that is a virus; and a vector that is a retrovirus) containing an isolated nucleic acid molecule encoding Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83.

The invention also features substantially pure Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 polypeptides; a substantially pure polypeptide which includes a first portion and a second portion, the first portion including a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide and the second portion including a detectable marker.

The invention also features an antibody that selectively binds to a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide (e.g., a monoclonal antibody).

The invention also features a pharmaceutical composition which includes a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide.

The invention includes methods for diagnosing a disorder associated with aberrant expression of a protein of the invention (i.e., Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83), the method including obtaining a biological sample from a patient and measuring the expression of the protein in the biological sample, wherein increased or decreased expression of the protein in the biological sample compared to a control indicates that the patient suffers from a disorder associated with aberrant expression of the protein.

The invention encompasses isolated nucleic acid molecules encoding Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 or a polypeptide fragment thereof; vectors containing these nucleic acid molecules; cells harboring recombinant DNA encoding Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83; fusion proteins which include all or a portion of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83; transgenic animals which express Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83; and recombinant knock-out animals which fail to express Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83.

The invention encompasses nucleic acids that have a sequence that is substantially identical to a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 nucleic acid sequence. A nucleic acid molecule which is substantially identical to a given reference nucleic acid molecule is hereby defined as a nucleic acid molecule having a sequence that has at least 85%, preferably 90%, and more

preferably 95%, 98%, 99% or more identity to the sequence of the given reference nucleic acid molecule.

The invention also includes polypeptides which are substantially identical to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 (e.g., polypeptides that are substantially identical to the polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10).

A polypeptide which is "substantially identical" to a given reference polypeptide molecule is a polypeptide having an amino acid sequence that has at least 85%, preferably 90%, and more preferably 95%, 98%, 99% or more identity to the amino acid sequence of the given reference polypeptide.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = # of identical positions/total # of positions (e.g., overlapping positions) x 100). Preferably, the two sequences are the same length.

The determination of percent homology between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two

sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. Id. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. *See* http://www.ncbi.nlm.nih.gov. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

The nucleic acid molecules of the invention can be inserted into vectors, as described below, which will facilitate expression of the insert. The nucleic acid molecules and the polypeptides they encode can be used directly as diagnostic or therapeutic agents, or (in the case of a polypeptide) can be used to generate antibodies that, in turn, are therapeutically useful. Accordingly, expression vectors containing the nucleic acid molecules of the invention, cells transfected with these vectors, the polypeptides expressed, and antibodies generated, against either the entire polypeptide or an antigenic fragment thereof, are among the preferred embodiments.

A transformed cell is any cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid encoding a polypeptide of the invention (e.g., a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide).

An isolated nucleic acid molecule is a nucleic acid molecule that is separated from the 5' and 3' coding sequences with which it is immediately contiguous in the naturally occurring genome of an organism. Isolated nucleic acid molecules include nucleic acid molecule which are not naturally occurring, e.g., nucleic acid molecules created by recombinant DNA techniques.

Nucleic acid molecules include both RNA and DNA, including cDNA, genomic DNA, and synthetic (e.g., chemically synthesized) DNA. Where single-stranded, the nucleic acid molecule may be a sense strand or an antisense strand.

The invention also encompasses nucleic acid molecules that hybridize, preferably under stringent conditions, to a nucleic acid molecule encoding a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide (e.g., a nucleic acid molecule having the sequence shown in SEQ ID NO:1, 3, 5, 7, or 9). Preferably the

hybridizing nucleic acid molecule consists of 400, more preferably 200 nucleotides. Preferred hybridizing nucleic acid molecules have a biological activity possessed by Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83.

The invention also features substantially pure or isolated Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptides, including those that correspond to various functional domains of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83, or fragments thereof.

The polypeptides of the invention can be produced recombinantly, chemically synthesized, or purified from tissues in which they are naturally expressed, according to standard biochemical methods of purification.

Also included in the invention are functional polypeptides, which possess one or more of the biological functions or activities of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. These functions include the ability to bind some or all of the proteins which normally bind to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. A functional polypeptide is also considered within the scope of the invention if it serves as an antigen for production of antibodies that specifically bind to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. In many cases, functional polypeptides retain one or more domains present in the naturally-occurring form of the polypeptide.

The functional polypeptides may contain a primary amino acid sequence that has been modified from those disclosed herein. Preferably these modifications consist of conservative amino acid substitutions, as described herein.

The terms "protein" and "polypeptide" are used herein to describe any chain of amino acids, regardless of length or post-translational modification (for

example, glycosylation or phosphorylation). Thus, the term "Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptides" includes full-length, naturally occurring Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 protein, as well a recombinantly or synthetically produced polypeptide that correspond to a full-length naturally occurring Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 protein or to particular domains or portions of a naturally occurring protein. The term also encompasses mature Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 which has an added amino-terminal methionine (useful for expression in prokaryotic cells).

The term "purified" as used herein refers to a nucleic acid or peptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized.

Polypeptides or other compounds of interest are said to be "substantially pure" when they are within preparations that are at least 60% by weight (dry weight) the compound of interest. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. Purity can be measured by any appropriate standard method, for example, by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

The invention also features antibodies, e.g., monoclonal, polyclonal, and engineered antibodies, which specifically bind Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. An antibody which specifically binds to a given antigen is an antibody that recognizes and binds to a particular antigen, but which does not substantially recognize or bind to other molecules in a sample, e.g., a

biological sample, which includes Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83.

The invention also features antagonists and agonists of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 that inhibit one or more of the biological activities of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. Suitable antagonists can include small molecules (i.e., molecules with a molecular weight below about 500), large molecules (i.e., molecules with a molecular weight above about 500), neutralizing antibodies, polypeptides which compete with a native form of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 for binding to a protein, and nucleic acid molecules that interfere with transcription of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 (for example, antisense nucleic acid molecules and ribozymes). Agonists of Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 also include small and large molecules, and antibodies other than neutralizing antibodies.

The invention also features molecules which can increase or decrease the expression of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 (e.g., by influencing transcription or translation). Small molecules (i.e., molecules with a molecular weight below about 500), large molecules (i.e., molecules with a molecular weight above about 500), and nucleic acid molecules that can be used to inhibit the expression of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 (for example, antisense and ribozyme molecules) or to enhance their expression (for example, molecules that bind to a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 transcription regulatory sequences and increase Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 transcription).

In addition, the invention features substantially pure polypeptides that functionally interact with Tango-

71, Tango-73, Tango-74, Tango-76, or Tango-83, and the nucleic acid molecules that encode them.

The invention encompasses methods for treating disorders associated with aberrant expression or activity of a protein of the invention (i.e., Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83). Thus, the invention includes methods for treating disorders associated with excessive expression or activity of a protein of the invention. Such methods entail administering a compound which decreases the expression or activity of the protein. The invention also includes methods for treating disorders associated with insufficient expression or activity of a protein of the invention. These methods entail administering a compound which increases the expression or activity of the protein.

The invention also features methods for detecting a protein of the invention (i.e., Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83). Such methods include: obtaining a biological sample; contacting the sample with an antibody that specifically binds the protein under conditions which permit specific binding; and detecting any antibody-protein complexes formed.

In addition, the present invention encompasses methods and compositions for the diagnostic evaluation, typing, and prognosis of disorders associated with inappropriate expression or activity of a protein of the invention. For example, the nucleic acid molecules of the invention can be used as diagnostic hybridization probes to detect, for example, inappropriate expression of a protein of the invention or mutations in the gene encoding a protein of the invention gene. Such methods may be used to classify cells by the level of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 expression.

The invention encompasses methods for diagnosing a disorder associated with aberrant activity of a protein

of the invention, the methods including obtaining a biological sample from a patient and measuring the activity of the protein in the biological sample, wherein increased or decreased activity in the biological sample compared to a control indicates that the patient suffers from a disorder associated with aberrant activity of the protein.

The nucleic acid molecules of the invention can be used as primers for diagnostic PCR analysis for the identification of gene mutations, allelic variations and regulatory defects in the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 gene. The present invention further provides for diagnostic kits for the practice of such methods.

The invention features methods of identifying compounds that modulate the expression or activity of a protein of the invention by assessing the expression or activity of the protein in the presence and absence of a selected compound. A difference in the level of expression or activity of the protein indicates that the selected compound is capable of modulating expression or activity of the protein. Expression can be assessed either at the level of gene expression (e.g., by measuring mRNA) or protein expression by techniques that are well known to skilled artisans. The activity of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 can be assessed functionally.

The preferred methods and materials are described below in examples which are meant to illustrate, not limit, the invention. Skilled artisans will recognize methods and materials that are similar or equivalent to those described herein, and that can be used in the practice or testing of the present invention.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as

commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described herein. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be limiting.

Other features and advantages of the invention will be apparent from the detailed description, and from the claims.

Brief Description of the Drawings

Figure 1 nucleotide sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) of Tango-71.

Figure 2 Nucleotide acid sequence (SEQ ID NO:3) and deduced amino acid sequence (SEQ ID NO:4) of Tango-73.

Figure 3 Nucleotide acid sequence (SEQ ID NO:5) and deduced amino acid sequence (SEQ ID NO:6) of Tango-74. The ATG encoding the first Met is boxed as is the ATC encoding the final Ile.

Figure 4 Nucleotide acid sequence of a 3' non-coding portion of Tango-74 (SEQ ID NO:11).

Figure 5 Alignment of a portion of the amino acid sequence of Tango-74 (SEQ ID NO:6) and the amino acid sequence of TRAIL.

Figure 6 Partial nucleotide sequence (SEQ ID NO:7) and deduced amino acid sequence (SEQ ID NO:8) of Tango-76.

Figure 7 Nucleotide sequence of a 5' portion of Tango-83 (SEQ ID NO:9).

Figure 8 Nucleotide sequence of a 3' portion of Tango-83 (SEQ ID NO:10).

Figure 9 Alignment of amino acid sequence of Tango-71 and the amino acid sequence of ADAMTS-1.

Figure 10 Alignment of the amino acid sequence of Tango-73 and the amino acid sequence of RVPI.

Figure 11 Alignment of the amino acid sequence of Tango-73 and TMVCF.

Figure 12 Northern blot analysis of Tango-73 mRNA.

Figure 13 Northern blot analysis of Tango-83 mRNA.

Figure 14 Alignment of amino acid sequence of Tango-76 and ADAMTS-1.

Detailed Description

Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 Nucleic Acid Molecules

The Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 nucleic acid molecules of the invention can be cDNA, genomic DNA, synthetic DNA, or RNA, and can be double-stranded or single-stranded (*i.e.*, either a sense or an antisense strand). Fragments of these molecules are also considered within the scope of the invention, and can be produced, for example, by the polymerase chain reaction (PCR) or generated by treatment with one or more restriction endonucleases. A ribonucleic acid (RNA) molecule can be produced by *in vitro* transcription.

The nucleic acid molecules of the invention can contain naturally occurring sequences, or sequences that differ from those that occur naturally, but, due to the degeneracy of the genetic code, encode the same polypeptide. In addition, these nucleic acid molecules are not limited to sequences that only encode

polypeptides, and thus, can include some or all of the non-coding sequences that lie upstream or downstream from a coding sequence.

The nucleic acid molecules of the invention can be synthesized (for example, by phosphoramidite-based synthesis) or obtained from a biological cell, such as the cell of a mammal. Thus, the nucleic acids can be those of a human, mouse, rat, guinea pig, cow, sheep, horse, pig, rabbit, monkey, dog, or cat. Combinations or modifications of the nucleotides within these types of nucleic acids are also encompassed.

In addition, the isolated nucleic acid molecules of the invention encompass fragments that are not found as such in the natural state. Thus, the invention encompasses recombinant molecules, such as those in which a nucleic acid molecule (for example, an isolated nucleic acid molecule encoding Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83) is incorporated into a vector (for example, a plasmid or viral vector) or into the genome of a heterologous cell (or the genome of a homologous cell, at a position other than the natural chromosomal location). Recombinant nucleic acid molecules and uses therefor are discussed further below.

In the event the nucleic acid molecules of the invention encode or act as antisense molecules, they can be used for example, to regulate translation of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 mRNA. Techniques associated with detection or regulation of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 expression are well known to skilled artisans and can be used to diagnose and/or treat disorders associated with aberrant Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 expression.

The invention also encompasses nucleic acid molecules that hybridize under stringent conditions to a

nucleic acid molecule encoding a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide (e.g., nucleic acid molecules having the sequence of the protein coding portion of SEQ ID NO:1, 3, 5, 7, or 9). The cDNA sequences described herein can be used to identify these hybridizing nucleic acids, which include, for example, nucleic acids that encode homologous polypeptides in other species and splice variants of the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 gene in humans or other mammals. Accordingly, the invention features methods of detecting and isolating these nucleic acid molecules. Using these methods, a sample (for example, a nucleic acid library, such as a cDNA or genomic library) is contacted (or "screened") with a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83-specific probe (for example, a fragment of SEQ ID NO:1, 3, 5, 7, or 9 that is at least 25 or 50 or 100 nucleotides long). The probe will selectively hybridize to nucleic acids encoding related polypeptides (or to complementary sequences thereof). The probe, which can contain at least 25 (for example, 25, 50, 100, or 200 nucleotides) can be produced using any of several standard methods (see, for example, Ausubel et al.,"Current Protocols in Molecular Biology, Vol. I," Green Publishing Associates, Inc., and John Wiley & Sons, Inc., NY, 1989). For example, the probe can be generated using PCR amplification methods in which oligonucleotide primers are used to amplify a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83-specific nucleic acid sequence that can be used as a probe to screen a nucleic acid library and thereby detect nucleic acid molecules (within the library) that hybridize to the probe.

One single-stranded nucleic acid is said to hybridize to another if a duplex forms between them. This occurs when one nucleic acid contains a sequence

that is the reverse and complement of the other (this same arrangement gives rise to the natural interaction between the sense and antisense strands of DNA in the genome and underlies the configuration of the "double helix"). Complete complementarity between the hybridizing regions is not required in order for a duplex to form; it is only necessary that the number of paired bases is sufficient to maintain the duplex under the hybridization conditions used.

Typically, hybridization conditions are of low to moderate stringency. These conditions favor specific interactions between completely complementary sequences, but allow some non-specific interaction between less than perfectly matched sequences to occur as well. After hybridization, the nucleic acids can be "washed" under moderate or high conditions of stringency to dissociate duplexes that are bound together by some non-specific interaction (the nucleic acids that form these duplexes are thus not completely complementary).

As is known in the art, the optimal conditions for washing are determined empirically, often by gradually increasing the stringency. The parameters that can be changed to affect stringency include, primarily, temperature and salt concentration. In general, the lower the salt concentration and the higher the temperature, the higher the stringency. Washing can be initiated at a low temperature (for example, room temperature) using a solution containing a salt concentration that is equivalent to or lower than that of the hybridization solution. Subsequent washing can be carried out using progressively warmer solutions having the same salt concentration. As alternatives, the salt concentration can be lowered and the temperature maintained in the washing step, or the salt concentration can be lowered and the temperature increased. Additional

parameters can also be altered. For example, use of a destabilizing agent, such as formamide, alters the stringency conditions.

In reactions where nucleic acids are hybridized, the conditions used to achieve a given level of stringency will vary. There is not one set of conditions, for example, that will allow duplexes to form between all nucleic acids that are 85% identical to one another; hybridization also depends on unique features of each nucleic acid. The length of the sequence, the composition of the sequence (for example, the content of purine-like nucleotides versus the content of pyrimidine-like nucleotides) and the type of nucleic acid (for example, DNA or RNA) affect hybridization. An additional consideration is whether one of the nucleic acids is immobilized (for example, on a filter).

An example of a progression from lower to higher stringency conditions is the following, where the salt content is given as the relative abundance of SSC (a salt solution containing sodium chloride and sodium citrate; 2X SSC is 10-fold more concentrated than 0.2X SSC). Nucleic acids are hybridized at 42°C in 2X SSC/0.1% SDS (sodium dodecylsulfate; a detergent) and then washed in 0.2X SSC/0.1% SDS at room temperature (for conditions of low stringency); 0.2X SSC/0.1% SDS at 42°C (for conditions of moderate stringency); and 0.1X SSC at 68°C (for conditions of high stringency). Washing can be carried out using only one of the conditions given, or each of the conditions can be used (for example, washing for 10-15 minutes each in the order listed above). Any or all of the washes can be repeated. As mentioned above, optimal conditions will vary and can be determined empirically.

A second set of conditions that are considered "stringent conditions" are those in which hybridization

is carried out at 50°C in Church buffer (7% SDS, 0.5% $NaHPO_4$, 1 M EDTA, 1% BSA) and washing is carried out at 50°C in 2X SSC.

Once detected, the nucleic acid molecules can be isolated by any of a number of standard techniques (see, for example, Sambrook et al., "Molecular Cloning, A Laboratory Manual," 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

The invention also encompasses: (a) expression vectors that contain any of the foregoing Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83-related coding sequences and/or their complements (that is, "antisense" sequence); (b) expression vectors that contain any of the foregoing Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83-related coding sequences operatively associated with a regulatory element (examples of which are given below) that directs the expression of the coding sequences; (c) expression vectors containing, in addition to sequences encoding a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide, nucleic acid sequences that are unrelated to nucleic acid sequences encoding Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83, such as molecules encoding a reporter or marker; and (d) genetically engineered host cells that contain any of the foregoing expression vectors and thereby express the nucleic acid molecules of the invention in the host cell.

Recombinant nucleic acid molecules can contain a sequence encoding a soluble Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide; mature Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83; or Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 having a signal sequence. A full length Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide; a domain of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83; or a fragment thereof may be fused to additional polypeptides, as

described below. Similarly, the nucleic acid molecules of the invention can encode the mature form of Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 or a form that encodes a polypeptide which facilitates secretion. In the latter instance, the polypeptide is typically referred to as a proprotein, which can be converted into an active form by removal of the signal sequence, for example, within the host cell. Proproteins can be converted into the active form of the protein by removal of the inactivating sequence.

The regulatory elements referred to above include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other elements, which are known to those skilled in the art, and which drive or otherwise regulate gene expression. Such regulatory elements include but are not limited to the cytomegalovirus hCMV immediate early gene, the early or late promoters of SV40 adenovirus, the lac system, the trp system, the TAC system, the TRC system, the major operator and promoter regions of phage A, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase, the promoters of acid phosphatase, and the promoters of the yeast α-mating factors.

Similarly, the nucleic acid can form part of a hybrid gene encoding additional polypeptide sequences, for example, sequences that function as a marker or reporter. Examples of marker or reporter genes include β-lactamase, chloramphenicol acetyltransferase (CAT), adenosine deaminase (ADA), aminoglycoside phosphotransferase (neo^r, $G418^r$), dihydrofolate reductase (DHFR), hygromycin-B-phosphotransferase (HPH), thymidine kinase (TK), lacZ (encoding β-galactosidase), and xanthine guanine phosphoribosyltransferase (XGPRT). As with many of the standard procedures associated with the

practice of the invention, skilled artisans will be aware of additional useful reagents, for example, of additional sequences that can serve the function of a marker or reporter. Generally, the hybrid polypeptide will include a first portion and a second portion; the first portion being a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide and the second portion being, for example, the reporter described above or an immunoglobulin constant region.

The expression systems that may be used for purposes of the invention include, but are not limited to, microorganisms such as bacteria (for example, *E. coli* and *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA, or cosmid DNA expression vectors containing the nucleic acid molecules of the invention; yeast (for example, *Saccharomyces* and *Pichia*) transformed with recombinant yeast expression vectors containing the nucleic acid molecules of the invention (preferably containing the nucleic acid sequence encoding Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83); insect cell systems infected with recombinant virus expression vectors (for example, baculovirus) containing the nucleic acid molecules of the invention; plant cell systems infected with recombinant virus expression vectors (for example, cauliflower mosaic virus (CaMV) and tobacco mosaic virus (TMV)) or transformed with recombinant plasmid expression vectors (for example, Ti plasmid) containing Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 nucleotide sequences; or mammalian cell systems (for example, COS, CHO, BHK, 293, VERO, HeLa, MDCK, WI38, and NIH 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (for example, the metallothionein promoter) or from mammalian viruses (for

example, the adenovirus late promoter and the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the gene product being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions containing Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptides or for raising antibodies to those polypeptides, vectors that are capable of directing the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited to, the *E. coli* expression vector pUR278 (Ruther et al., *EMBO J.* 2:1791, 1983), in which the coding sequence of the insert may be ligated individually into the vector in frame with the lacZ coding region so that a fusion protein is produced; pIN vectors (Inouye and Inouye, *Nucleic Acids Res.* 13:3101-3109, 1985; Van Heeke and Schuster, *J. Biol. Chem.* 264:5503-5509, 1989); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear polyhidrosis virus (AcNPV) can be used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The coding sequence of the insert may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus

and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed. (for example, see Smith et al., *J. Virol.* 46:584, 1983; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the nucleic acid molecule of the invention may be ligated to an adenovirus transcription/translation control complex, for example, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (for example, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 gene product in infected hosts (for example, see Logan and Shenk, *Proc. Natl. Acad. Sci. USA* 81:3655-3659, 1984). Specific initiation signals may also be required for efficient translation of inserted nucleic acid molecules. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire gene or cDNA, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the

initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., *Methods in Enzymol.* 153:516-544, 1987).

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (for example, glycosylation) and processing (for example, cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. The mammalian cell types listed above are among those that could serve as suitable host cells.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 sequences described above may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements

(for example, promoter, enhancer sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method can advantageously be used to engineer cell lines which express Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the gene product.

A number of selection systems can be used. For example, the herpes simplex virus thymidine kinase (Wigler, et al., *Cell* 11:223, 1977), hypoxanthine-guanine phosphoribosyltransferase (Szybalska and Szybalski, *Proc. Natl. Acad. Sci. USA* 48:2026, 1962), and adenine phosphoribosyltransferase (Lowy, et al., *Cell* 22:817, 1980) genes can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells, respectively. Also, anti-metabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler et al., *Proc. Natl. Acad. Sci. USA* 77:3567, 1980; O'Hare et al., *Proc. Natl. Acad. Sci. USA* 78:1527, 1981); gpt, which confers resistance to mycophenolic acid (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072, 1981); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin et al., *J. Mol. Biol.* 150:1, 1981); and hygro, which confers resistance to hygromycin (Santerre et al., *Gene* 30:147, 1984).

Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 nucleic acid molecules are useful in genetic mapping and chromosome identification.

Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 Polypeptides

The Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 polypeptides described herein are those encoded by any of the nucleic acid molecules described above and include Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 fragments, mutants, truncated forms, and fusion proteins. These polypeptides can be prepared for a variety of uses, including but not limited to the generation of antibodies, as reagents in diagnostic assays, for the identification of other cellular gene products or compounds that can modulate the activity or expression of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83, and as pharmaceutical reagents useful for the treatment of disorders associated with aberrant expression or activity of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83.

Preferred polypeptides are substantially pure Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 polypeptides, including those that correspond to the polypeptide with an intact signal sequence, the secreted form of a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide. Especially preferred are polypeptides that are soluble under normal physiological conditions.

The invention also encompasses polypeptides that are functionally equivalent to Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83. These polypeptides are equivalent to Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 in that they are capable of carrying out one or more of the functions of Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 in a biological system. Preferred

Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 polypeptides have 20%, 40%, 50%, 75%, 80%, or even 90% of one or more of the biological activities of the full-length, mature human form of Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83. Such comparisons are generally based on an assay of biological activity in which equal concentrations of the polypeptides are used and compared. The comparison can also be based on the amount of the polypeptide required to reach 50% of the maximal stimulation obtainable.

Functionally equivalent proteins can be those, for example, that contain additional or substituted amino acid residues. Substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. Amino acids that are typically considered to provide a conservative substitution for one another are specified in the summary of the invention.

Polypeptides that are functionally equivalent to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 can be made using random mutagenesis techniques well known to those skilled in the art. It is more likely, however, that such polypeptides will be generated by site-directed mutagenesis (again using techniques well known to those skilled in the art). These polypeptides may have increased functionality or decreased functionality.

To design functionally equivalent polypeptides, it is useful to distinguish between conserved positions and variable positions. This can be done by aligning the amino acid sequence of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 with the amino acid sequence of the homologons protein from another species. Skilled artisans will recognize that conserved amino acid residues are more likely to be necessary for preservation

of function. Thus, it is preferable that conserved residues are not altered.

Mutations within the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 coding sequence can be made to generate variant Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 genes that are better suited for expression in a selected host cell. For example, N-linked glycosylation sites can be altered or eliminated to achieve, for example, expression of a homogeneous product that is more easily recovered and purified from yeast hosts which are known to hyperglycosylate N-linked sites. To this end, a variety of amino acid substitutions at one or both of the first or third amino acid positions of any one or more of the glycosylation recognition sequences which occur, and/or an amino acid deletion at the second position of any one or more of such recognition sequences, will prevent glycosylation at the modified tripeptide sequence (see, for example, Miyajima et al., *EMBO J.* 5:1193, 1986).

Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (*Proc. Natl. Acad. Sci. USA* 88: 8972-8976, 1991). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni^{2+}·nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

The polypeptides of the invention can be expressed fused to another polypeptide, for example, a marker

polypeptide or fusion partner. For example, the polypeptide can be fused to a hexa-histidine tag to facilitate purification of bacterially expressed protein or a hemagglutinin tag to facilitate purification of protein expressed in eukaryotic cells.

The polypeptides of the invention can be chemically synthesized (for example, see Creighton, "Proteins: Structures and Molecular Principles," W.H. Freeman & Co., NY, 1983), or, perhaps more advantageously, produced by recombinant DNA technology as described herein. For additional guidance, skilled artisans may consult Ausubel et al. (*supra*), Sambrook et al. ("Molecular Cloning, A Laboratory Manual," Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989), and, particularly for examples of chemical synthesis Gait, M.J. Ed. ("Oligonucleotide Synthesis," IRL Press, Oxford, 1984).

The invention also features polypeptides that interact with Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 (and the genes that encode them) and thereby alter the function of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. Interacting polypeptides can be identified using methods known to those skilled in the art. One suitable method is the "two-hybrid system," which detects protein interactions *in vivo* (Chien et al., *Proc. Natl. Acad. Sci. USA*, 88:9578, 1991). A kit for practicing this method is available from Clontech (Palo Alto, CA).

Transgenic animals

Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 polypeptides can also be expressed in transgenic animals. These animals represent a model system for the study of disorders that are caused by or exacerbated by overexpression or underexpression of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83, and for the development

of therapeutic agents that modulate the expression or activity of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83.

Transgenic animals can be farm animals (pigs, goats, sheep, cows, horses, rabbits, and the like) rodents (such as rats, guinea pigs, and mice), non-human primates (for example, baboons, monkeys, and chimpanzees), and domestic animals (for example, dogs and cats). Transgenic mice are especially preferred.

Any technique known in the art can be used to introduce a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 transgene into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (U.S. Pat. No. 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten et al., *Proc. Natl. Acad. Sci., USA* 82:6148, 1985); gene targeting into embryonic stem cells (Thompson et al., *Cell* 56:313, 1989); and electroporation of embryos (Lo, *Mol. Cell. Biol.* 3:1803, 1983).

The present invention provides for transgenic animals that carry a the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 transgene in all their cells, as well as animals that carry a transgene in some, but not all of their cells. That is, the invention provides for mosaic animals. The transgene can be integrated as a single transgene or in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene can also be selectively introduced into and activated in a particular cell type (Lasko et al., *Proc. Natl. Acad. Sci. USA* 89:6232, 1992). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

When it is desired that the Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be used, vectors containing some nucleotide sequences homologous to an endogenous Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene also can be selectively introduced into a particular cell type, thus inactivating the endogenous Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 gene in only that cell type (Gu et al., *Science* 265:103, 1984). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. Once transgenic animals have been generated, expression of the recombinant Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 gene can be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to determine whether integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and RT-PCR. Samples of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 gene-expressing tissue can also be evaluated immunocytochemically using antibodies specific for the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 transgene product.

For a review of techniques that can be used to generate and assess transgenic animals, skilled artisans

can consult Gordon (*Intl. Rev. Cytol.* 115:171-229, 1989), and may obtain additional guidance from, for example: Hogan et al. "Manipulating the Mouse Embryo" (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1986; Krimpenfort et al., *Bio/Technology* 9:86, 1991; Palmiter et al., *Cell* 41:343, 1985; Kraemer et al., "Genetic Manipulation of the Early Mammalian Embryo," Cold Spring Harbor Press, Cold Spring Harbor, NY, 1985; Hammer et al., *Nature* 315:680, 1985; Purcel et al., *Science*, 244:1281, 1986; Wagner et al., U.S. Patent No. 5,175,385; and Krimpenfort et al., U.S. Patent No. 5,175,384 (the latter two publications are hereby incorporated by reference).

Anti-Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 Antibodies

Human Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 polypeptides (or immunogenic fragments or analogs) can be used to raise antibodies useful in the invention; such polypeptides can be produced by recombinant techniques or synthesized (see, for example, "Solid Phase Peptide Synthesis," *supra*; Ausubel et al., *supra*). In general, the peptides can be coupled to a carrier protein, such as KLH, as described in Ausubel et al., *supra*, mixed with an adjuvant, and injected into a host mammal. Antibodies can be purified by peptide antigen affinity chromatography.

In particular, various host animals can be immunized by injection with a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 protein or polypeptide. Host animals include rabbits, mice, guinea pigs, and rats. Various adjuvants that can be used to increase the immunological response depend on the host species and include Freund's adjuvant (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet

hemocyanin, and dinitrophenol. Potentially useful human adjuvants include BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. Polyclonal antibodies are heterogeneous populations of antibody molecules that are contained in the sera of the immunized animals.

Antibodies within the invention therefore include polyclonal antibodies and, in addition, monoclonal antibodies, humanized or chimeric antibodies, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, and molecules produced using a Fab expression library.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, can be prepared using the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptides described above and standard hybridoma technology (see, for example, Kohler et al., *Nature* 256:495, 1975; Kohler et al., *Eur. J. Immunol.* 6:511, 1976; Kohler et al., *Eur. J. Immunol.* 6:292, 1976; Hammerling et al., "Monoclonal Antibodies and T Cell Hybridomas," Elsevier, NY, 1981; Ausubel et al., *supra*).

In particular, monoclonal antibodies can be obtained by any technique that provides for the production of antibody molecules by continuous cell lines in culture such as described in Kohler et al., *Nature* 256:495, 1975, and U.S. Patent No. 4,376,110; the human B-cell hybridoma technique (Kosbor et al., *Immunology Today* 4:72, 1983; Cole et al., *Proc. Natl. Acad. Sci. USA* 80:2026, 1983), and the EBV-hybridoma technique (Cole et al., "Monoclonal Antibodies and Cancer Therapy," Alan R. Liss, Inc., pp. 77-96, 1983). Such antibodies can be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated *in vitro* or *in vivo*. The ability to produce high titers of mAbs *in vivo* makes this a particularly useful method of production.

Once produced, polyclonal or monoclonal antibodies are tested for specific Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 recognition by Western blot or immunoprecipitation analysis by standard methods, e.g., as described in Ausubel et al., *supra*. Antibodies that specifically recognize and bind to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 are useful. For example, such antibodies can be used in an immunoassay to monitor the level of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 produced by a mammal (for example, to determine the amount or subcellular location of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83).

Preferably, antibodies of the invention are produced using fragments of the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 protein which lie outside highly conserved regions and appear likely to be antigenic, by criteria such as high frequency of charged residues. In one specific example, such fragments are generated by standard techniques of PCR, and are then cloned into the pGEX expression vector (Ausubel et al., *supra*). Fusion proteins are expressed in *E. coli* and purified using a glutathione agarose affinity matrix as described in Ausubel, et al., *supra*.

In some cases it may be desirable to minimize the potential problems of low affinity or specificity of antisera. In such circumstances, two or three fusions can be generated for each protein, and each fusion can be injected into at least two rabbits. Antisera can be raised by injections in a series, preferably including at least three booster injections.

Antisera is also checked for its ability to immunoprecipitate recombinant Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 protein or control proteins, such as glucocorticoid receptor, CAT, or luciferase.

The antibodies can be used, for example, in the detection of the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 in a biological sample as part of a diagnostic assay. Antibodies also can be used in a screening assay to measure the effect of a candidate compound on expression or localization of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. Additionally, antibodies can be used in conjunction with the gene therapy techniques described to, for example, evaluate the normal and/or engineered Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83-expressing cells prior to their introduction into the patient. Such antibodies additionally can be used in a method for inhibiting abnormal Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 activity.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., *Proc. Natl. Acad. Sci. USA*, 81:6851, 1984; Neuberger et al., *Nature*, 312:604, 1984; Takeda et al., *Nature*, 314:452, 1984) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent Nos. 4,946,778, 4,946,778, and 4,704,692) can be adapted to produce single chain antibodies against a Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments that recognize and bind to specific epitopes can be generated by known techniques. For example, such fragments include but are not limited to $F(ab')_2$ fragments that can be produced by pepsin digestion of the antibody molecule, and Fab fragments that can be generated by reducing the disulfide bridges of $F(ab')_2$ fragments. Alternatively, Fab expression libraries can be constructed (Huse et al., *Science*, 246:1275, 1989) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

Antibodies to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 can, in turn, be used to generate anti-idiotype antibodies that resemble a portion of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 using techniques well known to those skilled in the art (see, e.g., Greenspan et al., *FASEB J.* 7:437, 1993; Nissinoff, *J. Immunol.* 147:2429, 1991). For example, antibodies that bind to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 and competitively inhibit the binding of a binding partner of the protein can be used to generate anti-idiotypes that resemble a binding partner binding domain of the protein and, therefore, bind and neutralize a binding partner of the protein. Such neutralizing anti-idiotypic antibodies or Fab fragments of such anti-idiotypic antibodies can be used in therapeutic regimens.

Antibodies can be humanized by methods known in the art. For example, monoclonal antibodies with a desired binding specificity can be commercially humanized (Scotgene, Scotland; Oxford Molecular, Palo Alto, CA). Fully human antibodies, such as those expressed in transgenic animals are also features of the invention (Green et al., Nature Genetics 7:13-21, 1994; see also U.S. Patents 5,545,806 and 5,569,825, both of which are hereby incorporated by reference).

The methods described herein in which anti-Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 antibodies are employed may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one specific Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 antibody reagent described herein, which may be conveniently used, for example, in clinical settings, to diagnose patients exhibiting symptoms of the disorders described below.

Antisense Nucleic Acids

Treatment regimes based on an "antisense" approach involve the design of oligonucleotides (either DNA or RNA) that are complementary to Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 mRNA. These oligonucleotides bind to the complementary Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required. A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarily to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarily and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it may contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the message, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work

most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs recently have been shown to be effective at inhibiting translation of mRNAs as well (Wagner, *Nature* 372:333, 1984). Thus, oligonucleotides complementary to either the 5' or 3' non-translated, non-coding regions of the mRNA, could be used in an antisense approach to inhibit translation of endogenous Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon.

Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the 5', 3', or coding region of mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides, or at least 50 nucleotides.

Regardless of the choice of target sequence, it is preferred that *in vitro* studies are first performed to quantitate the ability of the antisense oligonucleotide to inhibit gene expression. It is preferred that these studies utilize controls that distinguish between antisense gene inhibition and nonspecific biological effects of oligonucleotides. It is also preferred that these studies compare levels of the target RNA or protein with that of an internal control RNA or protein. Additionally, it is envisioned that results obtained using the antisense oligonucleotide are compared with those obtained using a control oligonucleotide. It is preferred that the control oligonucleotide is of

approximately the same length as the test oligonucleotide and that the nucleotide sequence of the oligonucleotide differs from the antisense sequence no more than is necessary to prevent specific hybridization to the target sequence.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (as described, e.g., in Letsinger et al., *Proc. Natl. Acad. Sci. USA* 86:6553, 1989; Lemaitre et al., *Proc. Natl. Acad. Sci. USA* 84:648, 1987; PCT Publication No. WO 88/09810) or the blood-brain barrier (see, for example, PCT Publication No. WO 89/10134), or hybridization-triggered cleavage agents (see, for example, Krol et al., *BioTechniques* 6:958, 1988), or intercalating agents (see, for example, Zon, *Pharm. Res.* 5:539, 1988). To this end, the oligonucleotide can be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethyl-aminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine,

1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-theouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 2-(3-amino-3-N-2-carboxypropl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal, or an analog of any of these backbones.

In yet another embodiment, the antisense oligonucleotide is an α-anomeric oligonucleotide. An α-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gautier et al., *Nucl. Acids. Res.* 15:6625, 1987). The oligonucleotide is a 2'-0-methylribonucleotide (Inoue et al., *Nucl. Acids Res.* 15:6131, 1987), or a chimeric RNA-DNA analog (Inoue et al., *FEBS Lett.* 215:327, 1987).

Antisense oligonucleotides of the invention can be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides can be synthesized by the method of Stein et al. (*Nucl. Acids Res.* 16:3209, 1988), and methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., *Proc. Natl. Acad. Sci. USA* 85:7448, 1988).

The antisense molecules should be delivered to cells that express the protein of interest *in vivo*. A number of methods have been developed for delivering antisense DNA or RNA to cells; e.g., antisense molecules can be injected directly into the tissue site, or modified antisense molecules, designed to target the desired cells (e.g., antisense linked to peptides or antibodies that specifically bind receptors or antigens expressed on the target cell surface) can be administered systemically.

However, it is often difficult to achieve intracellular concentrations of the antisense molecule sufficient to suppress translation of endogenous mRNAs. Therefore, a preferred approach uses a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong *pol* III or *pol* II promoter. The use of such a construct to transfect target cells in the patient will result in the transcription of sufficient amounts of single stranded RNAs that will form complementary base pairs with the endogenous Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 transcripts and thereby prevent translation. For example, a vector can be introduced *in vivo* such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or

become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA.

Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to: the SV40 early promoter region (Bernoist et al., *Nature* 290:304, 1981); the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., *Cell* 22:787-797, 1988); the herpes thymidine kinase promoter (Wagner et al., *Proc. Natl. Acad. Sci. USA* 78:1441, 1981); or the regulatory sequences of the metallothionein gene (Brinster et al., *Nature* 296:39, 1988).

Ribozymes

Ribozyme molecules designed to catalytically cleave Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 mRNA transcripts also can be used to prevent translation and expression of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 (see, e.g., PCT Publication WO 90/11364; Saraver et al., *Science* 247:1222, 1990). While various ribozymes that cleave mRNA at site-specific recognition sequences can be used to destroy Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art (Haseloff et al., *Nature*

334:585, 1988). There are numerous examples of potential hammerhead ribozyme cleavage sites within the nucleotide sequence of human Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 cDNA. Preferably, the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the mRNA, i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes"), such as the one that occurs naturally in *Tetrahymena Thermophila* (known as the IVS or L-19 IVS RNA), and which has been extensively described by Cech and his collaborators (Zaug et al., *Science* 224:574, 1984; Zaug et al., *Science*, 231:470, 1986; Zug et al., *Nature* 324:429, 1986; PCT Application No. WO 88/04300; and Been et al., *Cell* 47:207, 1986). The Cech-type ribozymes have an eight base-pair sequence that hybridizes to a target RNA sequence, whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes that target eight base-pair active site sequences present in Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.), and should be delivered to cells which express Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 *in vivo*. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive *pol* III or *pol* II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous messages and inhibit translation. Because ribozymes, unlike antisense molecules, are catalytic, a

lower intracellular concentration is required for efficiency.

Other Methods for Reducing Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 Expression

Endogenous Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 gene expression can also be reduced by inactivating the endogenous gene or its promoter using targeted homologous recombination (see, e.g., U.S. Patent No. 5,464,764). For example, a mutant, non-functional Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous gene (either the coding regions or regulatory regions) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express the endogenous gene *in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 gene. Such approaches are particularly suited for use in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83. However, this approach can be adapted for use in humans, provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors.

Alternatively, endogenous Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the gene (*i.e.*, the promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells in the body (Helene *Anticancer Drug Res.*

6:569, 1981; Helene et al., *Ann. N.Y. Acad. Sci.* 660:27, 1992; and Maher, *Bioassays* 14:807, 1992).

Detecting Proteins Associated with Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83

The invention also features polypeptides which interact with Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. Any method suitable for detecting protein-protein interactions may be employed for identifying transmembrane proteins, intracellular, or extracellular proteins that interact with Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. Among the traditional methods which may be employed are co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns of cell lysates or proteins obtained from cell lysates and the use of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 to identify proteins in the lysate that interact with Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. For these assays, the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypetide can be: a full length Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83; a soluble extracellular domain of Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83; or some other suitable Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide. Once isolated, such an interacting protein can be identified and cloned and then used, in conjunction with standard techniques, to identify proteins with which it interacts. For example, at least a portion of the amino acid sequence of a protein which interacts with the Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 can be ascertained using techniques well known to those of skill in the art, such as via the Edman degradation technique. The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for gene sequences encoding

the interacting protein. Screening may be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well-known (Ausubel, *supra;* and "PCR Protocols: A Guide to Methods and Applications," Innis et al., eds. Academic Press, Inc., NY, 1990).

Additionally, methods may be employed which result directly in the identification of genes which encode proteins which interact with Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83. These methods include, for example, screening expression libraries, in a manner similar to the well known technique of antibody probing of λgt11 libraries, using labeled Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 polypeptide or a Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 fusion protein, e.g., a Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 polypeptide or domain fused to a marker such as an enzyme, fluorescent dye, a luminescent protein, or to an IgFc domain.

There are also methods which are capable of detecting protein interaction. A method which detects protein interactions *in vivo* is the two-hybrid system (Chien et al., *Proc. Natl. Acad. Sci. USA,* 88:9578, 1991). A kit for practicing this method is available from Clontech (Palo Alto, CA).

Compounds which bind Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83

Compounds which bind Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 can be identified using any standard binding assay. For example, candidate compounds can be bound to a solid support. Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83 is then exposed to the immobilized compound and binding is measured (European Patent Application 84/03564).

EXAMPLES

Tango-71 cDNA (Fig. 1; SEQ ID NO:1) was isolated from human melanocytes as follows.

Human melanocytes (Clonetics Corporation; San Diego, CA) were expanded in culture with Melanocyte Growth Media (MGM; Clonetics) according to the recommendations of the supplier. When the cells reached ~80-90% confluence, they were starved in MGM without growth factors for 46 hours. The starved cells were then stimulated with complete MGM supplemented with 20 ng/ml TNF (Gibco BRL; Gaithersburg, MD) and cycloheximide (CHI;40 micrograms/ml) for 4 hours. Total RNA was isolated using the RNeasy Midi Kit (Qiagen; Chatsworth, CA), and the poly A+ fraction was further purified using Oligotex beads (Qiagen).

Three micrograms of poly A+ RNA were used to synthesize a cDNA library using the Superscript cDNA Synthesis kit (Gibco BRL). Complementary DNA was directionally cloned into the expression plasmid pMET7 using the SalI and NotI sites in the polylinker to construct a plasmid library. Transformants were picked and grown up for single-pass sequencing. Additionally, astrocyte cDNA was ligated into the SalI/NotI sites of the ZipLox vector (Gibco BRL) for construction of a lambda phage cDNA library.

Northern blot analysis of Tango-71 expression was performed using Tango-71 labeled with ^{32}P-dCTP using the Prime-It kit (Stratagene, LaJolla, CA). Human mRNA blots (MTNI and MTNII; Clonetech; Palo Alto, CA) were probed and washed at high stringency as recommended by the manufacturer. Tango-71 is expressed as an approximately 6.0 kb transcript in all tissues: heart brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testes, ovary, small intestine, colon, PBLs.

The amino acid sequence of a portion of Tango-71 is 90% identical to the amino acid sequence of murine ADAMTS-1 (FIG. 9), a cellular disintegrin and metalloprotease that is thought to be involved in inflammation and development of cancer cachexia (Kuno et al., *J. Biol. Chem.* 272:556, 1997). Based on sequence comparison to ADAMTS-1, Tango-71, using the amino acid numbering in Figure 9, has the following domains: amino acids 1-160 (metalloproteinase domain, partial); amino acids 170-242 (disintegrin domain); amino acids 257-307 (thrombospondin domain). A less apparent thrombospondin domain is present at amino acid 558-608. Portions of Tango-71 shown in Figure 1, but not in Figure 9, may also be homologous ADAMTS-1. Tango-71 may represent the human homolog of ADAMTS-1 or a splice variant thereof.

Tango-71 expression may be androgen regulated. Tango-71 expression in LNCaP cells, an androgen-dependent prostate cancer cell line, is induced by R1881, a testosterone analog. Tango-71 expression is downregulated in LNCaP cells treated with casodex, an anti-androgen.

Tango-73 cDNA (Fig. 2; SEQ ID NO:2) was isolated from human prostate epithelial cells as follows.

Human prostate epithelial cells (Clonetics) were expanded in culture with Prostate Epithelial Growth Medium (PEGM) (Clonetics). When cells reached confluence cells were grown in Prostate Basal Media (Clonetics) for 24 hours. They were stimulated with PEGM (prostate epithelial growth medium; Clonetics) and 40 ug/ml cycloheximide for 3 hours.

Total RNA was isolated using the RNeasy Midi Kit (Qiagen). Poly (A)+ was isolated using the Oligotex beads (Qiagen). Next, cDNA was constructed using the Superscript-cDNA Synthesis Kit (Gibco BRL). The cDNA was cloned into the expression vector pMET7 using the SalI

and NotI sites in the polylinker. Transformants were picked and sequenced.

Northern blot analysis of Tango-73 expression was carried out as described above. This analysis revealed the presence of 4.0 kb and 3.0 kb transcripts in the placenta and liver. A 4.0 kb transcript was present in lung, kidney, thymus, prostate, spleen, testes, and colon, with the highest expression in lung, pancreas, prostate, and testes.

The amino acid sequence of Tango-73 is 48% identical to rat RVP.1 (Briehl et al., *Mol. Endocrinol.* 5:1381, 1991) and 46.1% identical to TMVCF (Sirotkin et al., *Genomics* 42:245, 1997).

RVP.1 is up-regulated during apoptosis (Briehl et al., *supra*). TMVCF, a 219 amino acid protein with two putative membrane spanning domains, is deleted in velo-cardio-facial syndrome (Sirotkin et al., *supra*).

Tango-83 (Figs. 7 and 8) and Tango-74 cDNAs (Fig. 5) were isolated from human astrocytes as follows.

Human astrocytes (Clonetics) were expanded in culture with Astocyte Growth Media (AGM; Clonetics) according to the recommendations of the supplier. When the cells reached ~80-90% confluence, they were stimulated with 200 units/ml Interleukin 1-Beta (Boehringer Mannheim) and cycloheximide (CHI: 40 micrograms/ml) for 4 hours. Total RNA was isolated using the RNeasy Midi Kit (Qiagen), and the poly A+ fraction was further purified using Oligotex beads (Qiagen).

Three micrograms of poly A+ RNA were used to synthesize a cDNA library using the Superscript cDNA Synthesis kit (Gibco BRL). Complementary DNA was directionally cloned into the expression plasmid pMET7 using the SalI and NotI sites in the polylinker to construct ā plasmid library. Transformants were picked and grown up for single-pass sequencing. Additionally,

astrocyte cDNA was ligated into the SalI/NotI sites of the ZipLox vector (Gibco BRL) for construction of a lambda phage cDNA library.

Northern blot analysis of Tango-83 expression, performed as described above, revealed that Tango-83 is expressed as an approximately 9.0 kb transcript in brain (FIG. 13).

Northern blot analysis, performed as described above, revealed that Tango-74 is expressed as an approximately 4.0 kb transcript in heart, brain, lung, liver, kidney, pancreas, spleen, prostate, testes, ovary, small intestine, colon and peripheral blood lymphocytes. Higher expression was seen in lung, liver, skeletal muscle, spleen, testes, colon and peripheral blood lymphocytes.

The amino acid sequence of Tango-74 is homologous to the amino acid sequence of the TRAIL receptor (Pan et al., *Science* 276:111, 1997) (FIG. 5).

Tango-76 cDNA (SEQ ID NO:7) was isolated form an adult rat frontal cortex library. The amino acid sequence of Tango-76 is homologous to the amino acid sequence of ADAMTS-1 (FIG. 14).

Northern blot analysis of human mRNA probed with a Tango-76 probe revealed a 4.2 kb band in lung. Analysis of rat mRNA revealed a weak 3.8 kb transcript in heart, brain, spleen, liver, skeletal muscle, and kidney and a weak 1.8 kb transcript in spleen and liver.

Effective Dose

Toxicity and therapeutic efficacy of the polypeptides of the invention and the compounds that modulate their expression or activity can be determined by standard pharmaceutical procedures, using either cells in culture or experimental animals to determine the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the

population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Polypeptides or other compounds that exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (that is, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

Formulations and Use

Pharmaceutical compositions for use in accordance with the present invention may be formulated in conventional manner using one or more physiologically acceptable carriers or excipients.

Thus, the compounds and their physiologically acceptable salts and solvates may be formulated for administration by inhalation or insufflation (either

through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (for example, pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (for example, lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (for example, magnesium stearate, talc or silica); disintegrants (for example, potato starch or sodium starch glycolate); or wetting agents (for example, sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (for example, sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (for example, lecithin or acacia); non-aqueous vehicles (for example, almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (for example, methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer salts, flavoring, coloring and sweetening agents as appropriate. Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

For buccal administration the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, for example, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, for example, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, for example, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, for example, in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, for example, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, for example, containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with

suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration.

The therapeutic compositions of the invention can also contain a carrier or excipient, many of which are known to skilled artisans. Excipients which can be used include buffers (for example, citrate buffer, phosphate buffer, acetate buffer, and bicarbonate buffer), amino acids, urea, alcohols, ascorbic acid, phospholipids, proteins (for example, serum albumin), EDTA, sodium chloride, liposomes, mannitol, sorbitol, and glycerol. The nucleic acids, polypeptides, antibodies, or modulatory compounds of the invention can be administered by any standard route of administration. For example, administration can be parenteral, intravenous, subcutaneous, intramuscular, intracranial, intraorbital, opthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, transmucosal, or oral. The modulatory compound can be formulated in various ways, according to the corresponding route of administration. For example, liquid solutions can be made for ingestion or injection; gels or powders can be made for ingestion, inhalation, or topical application. Methods for making such formulations are well known and can be found in, for example, "Remington's Pharmaceutical Sciences." It is expected that the preferred route of administration will be intravenous.

What is claimed is:

1. An isolated nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof;

b) a nucleic acid molecule comprising a fragment of at least 300 nucleotides of the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof;

c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____;

d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or the polypeptide encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____; and

e) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the nucleic

acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, or a complement thereof under stringent conditions.

2. The isolated nucleic acid molecule of claim 1, which is selected from the group consisting of:

a) a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, or a complement thereof; and

b) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____.

3. The nucleic acid molecule of claim 1 further comprising vector nucleic acid sequences.

4. The nucleic acid molecule of claim 1 further comprising nucleic acid sequences encoding a heterologous polypeptide.

5. A host cell which contains the nucleic acid molecule of claim 1.

6. The host cell of claim 5 which is a mammalian host cell.

7. A non-human mammalian host cell containing the nucleic acid molecule of claim 1.

8. An isolated polypeptide selected from the group consisting of:

a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4, SEQ ID NO:6, or SEQ ID NO:8, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2 or SEQ ID NO:4, SEQ ID NO:6, or SEQ ID NO:8;

b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, or a complement thereof under stringent conditions; and

c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, or a complement thereof.

9. The isolated polypeptide of claim 8 comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____.

10. The polypeptide of claim 8 further comprising heterologous amino acid sequences.

11. An antibody which selectively binds to a polypeptide of claim 8.

12. A method for producing a polypeptide selected from the group consisting of:

a) a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____;

b) a polypeptide comprising a fragment of the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____; and

c) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:10, or an amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, or a complement thereof under stringent conditions;

comprising culturing the host cell of claim 5 under conditions in which the nucleic acid molecule is expressed.

13. A method for detecting the presence of a polypeptide of claim 8 in a sample, comprising:

a) contacting the sample with a compound which selectively binds to a polypeptide of claim 8; and

b) determining whether the compound binds to the polypeptide in the sample.

14. The method of claim 13, wherein the compound which binds to the polypeptide is an antibody.

15. A kit comprising a compound which selectively binds to a polypeptide of claim 8 and instructions for use.

16. A method for detecting the presence of a nucleic acid molecule of claim 1 in a sample, comprising the steps of:

a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample.

17. The method of claim 16, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

18. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of claim 1 and instructions for use.

19. A method for identifying a compound which binds to a polypeptide of claim 8 comprising the steps of:

a) contacting a polypeptide, or a cell expressing a polypeptide of claim 8 with a test compound; and

b) determining whether the polypeptide binds to the test compound.

20. The method of claim 19, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

a) detection of binding by direct detecting of test compound/polypeptide binding;

b) detection of binding using a competition binding assay;

c) detection of binding using an assay for Tango-71, Tango-73, Tango-74, Tango-76, or Tango-83-mediated signal transduction.

21. A method for modulating the activity of a polypeptide of claim 8 comprising contacting a polypeptide or a cell expressing a polypeptide of claim 8 with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

22. A method for identifying a compound which modulates the activity of a polypeptide of claim 8, comprising:

a) contacting a polypeptide of claim 8 with a test compound; and

b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

```
GTCGACCCACGCGTCCGAGCGGCTCCGAGCCAGGGGCTATTGCAAAGCCAGGGTGCGCTACCGGACGGAGAGGGGAGAG  79

CCCTGAGCAGAGTGAGCAACATCGCAGCCAAGGCGGAGGCCGAAGAGGGGCGCCAGGCACCAATCTCCGCGTTGCCTCA 158

GCCCCGGAGGCGCCCCAGAGCGCTTCTTGTCCCAGCAGAGCCACTCTGCCTGCGCCTGCCTCTCAGTGTCTCCAACTTT 237

GCGCTGGAAGAAAAACTTCCCGCGCGCCGGCAGAACTGCAGCGCCTCCTCTTAGTGACTCCGGGAGCTTCGGCTGTAGC 316

CGGCTCTGCGCGCCCTTCCAACGAATAATAGAAATTGTTAATTTTAACAATCCAGAGCAGGCCAACGAGGCTTTGCTCT 395

                                                                M  Q  R          3
CCCGACCCGAACTAAAGCTCCCTCGCTCCGTGCGCTGCTACGAACGGTGTCTCCTGGGGCTCCA ATG CAG CGA     468

 A  V  P  E  G  F  G  R  R  K  L  G  S  D  M  G  N  A  E  R         23
GCT GTG CCC GAG GGG TTC GGA AGG CGC AAG CTG GGC AGC GAC ATG GGG AAC GCG GAG CGG   528

 A  P  G  S  R  S  F  G  P  V  P  T  L  L  L  L  A  A  A  L         43
GCT CCG GGG TCT CGG AGC TTT GGG CCC GTA CCC ACG CTG CTG CTG CTC GCC GCG GCG CTA   588

 L  A  V  S  D  A  L  G  R  P  S  E  E  D  E  E  L  V  V  P         63
CTG GCC GTG TCG GAC GCA CTC GGG CGC CCC TCC GAG GAG GAC GAG GAG CTA GTG GTG CCG   648

 E  L  E  R  A  P  G  H  G  T  T  R  L  R  L  H  A  F  D  Q         83
GAG CTG GAG CGC GCC CCG GGA CAC GGG ACC ACG CGC CTC CGC CTG CAC GCC TTT GAC CAG   708

 Q  L  D  L  E  L  R  P  D  S  S  F  L  A  P  G  F  T  L  Q        103
CAG CTG GAT CTG GAG CTG CGG CCC GAC AGC AGC TTT TTG GCG CCC GGC TTC ACG CTC CAG   768

 N  V  G  R  K  S  G  S  E  T  P  L  P  E  T  D  L  A  H  C        123
AAC GTG GGG CGC AAA TCC GGG TCC GAG ACG CCG CTT CCG GAA ACC GAC CTG GCG CAC TGC   828

 F  Y  S  G  T  V  N  G  D  P  S  S  A  A  A  L  S  L  C  E        143
TTC TAC TCC GGC ACC GTG AAT GGC GAT CCC AGC TCG GCT GCC GCC CTC AGC CTC TGC GAG   888

 G  V  R  G  A  F  Y  L  L  G  E  A  Y  F  I  Q  P  L  P  A        163
GGC GTG CGC GGC GCC TTC TAC CTG CTG GGG GAG GCG TAT TTC ATC CAG CCG CTG CCC GCC   948

 A  S  E  R  L  A  T  A  A  P  G  E  K  P  P  A  P  L  Q  F        183
GCC AGC GAG CGC CTC GCC ACC GCC GCC CCA GGG GAG AAG CCG CCG GCA CCA CTA CAG TTC  1008

 H  L  L  R  R  N  R  Q  G  D  V  G  G  T  C  G  V  V  D  D        203
CAC CTC CTG CGG CGG AAT CGG CAG GGC GAC GTA GGC GGC ACG TGC GGG GTC GTG GAC GAC  1068

 E  P  R  P  T  G  K  A  E  T  E  D  E  D  E  G  T  E  G  E        223
GAG CCC CGG CCG ACT GGG AAA GCG GAG ACC GAA GAC GAG GAC GAA GGG ACT GAG GGC GAG  1128

 D  E  G  P  Q  W  S  P  Q  D  P  A  L  Q  G  V  G  Q  P  T        243
GAC GAA GGG CCT CAG TGG TCG CCG CAG GAC CCG GCA CTG CAA GGC GTA GGA CAG CCC ACA  1188

 G  T  G  S  I  R  K  K  R  F  V  S  S  H  R  Y  V  E  T  M        263
GGA ACT GGA AGC ATA AGA AAG AAG CGA TTT GTG TCC AGT CAC CGC TAT GTG GAA ACC ATG  1248

 L  V  A  D  Q  S  M  A  E  F  H  G  S  G  L  K  H  Y  L  L        283
CTT GTG GCA GAC CAG TCG ATG GCA GAA TTC CAC GGC AGT GGT CTA AAG CAT TAC CTT CTC  1308
```

FIGURE 1 (1 of 4)

```
T  L  P  S  V  A  A  R  L  Y  K  H  P  S  I  R  N  S  V  S   303
ACG TTG TTT TCG GTG GCA GCC AGA TTG TAC AAA CAC CCC AGC ATT CGT AAT TCA GTT AGC  1368

L  V  V  V  K  I  L  V  I  H  D  E  Q  K  G  P  E  V  T  S   323
CTG GTG GTG GTG AAG ATC TTG GTC ATC CAC GAT GAA CAG AAG GGG CCG GAA GTG ACC TCC  1428

N  A  A  L  T  L  R  N  F  C  N  W  Q  K  Q  H  N  P  P  S   343
AAT GCT GCC CTC ACT CTG CGG AAC TTT TGC AAC TGG CAG AAG CAG CAC AAC CCA CCC AGT  1488

D  R  D  A  E  H  Y  D  T  A  I  L  F  T  R  Q  D  L  C  G   363
GAC CGG GAT GCA GAG CAC TAT GAC ACA GCA ATT CTT TTC ACC AGA CAG GAC TTG TGT GGG  1548

S  Q  T  C  D  T  L  G  M  A  D  V  G  T  V  C  D  P  S  R   383
TCC CAG ACA TGT GAT ACT CTT GGG ATG GCT GAT GTT GGA ACT GTG TGT GAT CCG AGC AGA  1608

S  C  S  V  I  E  D  D  G  L  Q  A  A  F  T  T  A  H  E  L   403
AGC TGC TCC GTC ATA GAA GAT GAT GGT TTA CAA GCT GCC TTC ACC ACA GCC CAT GAA TTA  1668

G  H  V  F  N  M  P  H  D  D  A  K  Q  C  A  S  L  N  G  V   423
GGC CAC GTG TTT AAC ATG CCA CAT GAT GAT GCA AAG CAG TGT GCC AGC CTT AAT GGT GTG  1728

N  Q  D  S  H  M  M  A  S  M  L  S  N  L  D  H  S  Q  P  W   443
AAC CAG GAT TCC CAC ATG ATG GCG TCA ATG CTT TCC AAC CTG GAC CAC AGC CAG CCT TGG  1788

S  P  C  S  A  Y  M  I  T  S  F  L  D  N  G  H  G  E  C  L   463
TCT CCT TGC AGT GCC TAC ATG ATT ACA TCA TTT CTG GAT AAT GGT CAT GGG GAA TGT TTG  1848

M  D  K  P  Q  N  P  I  Q  L  P  G  D  L  P  G  T  S  Y  D   483
ATG GAC AAG CCT CAG AAT CCC ATA CAG CTC CCA GGC GAT CTC CCT GGC ACC TCG TAC GAT  1908

A  N  R  Q  C  Q  F  T  F  G  E  D  S  K  H  C  P  D  A  A   503
GCC AAC CGG CAG TGC CAG TTT ACA TTT GGG GAG GAC TCC AAA CAC TGC CCC GAT GCA GCC  1968

S  T  C  S  T  L  W  C  T  G  T  S  G  G  V  L  V  C  Q  T   523
AGC ACA TGT AGC ACC TTG TGG TGT ACC GGC ACC TCT GGT GGG GTG CTG GTG TGT CAA ACC  2028

K  H  F  P  W  A  D  G  T  S  C  G  E  G  K  W  C  I  N  G   543
AAA CAC TTC CCG TGG GCG GAT GGC ACC AGC TGT GGA GAA GGG AAA TGG TGT ATC AAC GGC  2088

K  C  V  N  K  T  D  R  K  H  F  D  T  P  F  H  G  S  W  G   563
AAG TGT GTG AAC AAA ACC GAC AGA AAG CAT TTT GAT ACG CCT TTT CAT GGA AGC TGG GGA  2148

M  W  G  P  W  G  D  C  S  R  T  C  G  G  G  V  Q  Y  T  M   583
ATG TGG GGG CCT TGG GGA GAC TGT TCG AGA ACG TGC GGT GGA GGA GTC CAG TAC ACG ATG  2208

R  E  C  D  N  P  V  P  K  N  G  G  K  Y  C  E  G  K  R  V   603
AGG GAA TGT GAC AAC CCA GTC CCA AAG AAT GGA GGG AAG TAC TGT GAA GGC AAA CGA GTG  2268

R  Y  R  S  C  N  L  E  D  C  P  D  N  N  G  K  T  F  R  E   623
CGC TAC AGA TCC TGT AAC CTT GAG GAC TGT CCA GAC AAT AAT GGA AAA ACC TTT AGA GAG  2328

E  Q  C  E  A  H  N  E  F  S  K  A  S  F  G  S  G  P  A  V   643
GAA CAA TGT GAA GCA CAC AAC GAG TTT TCA AAA GCT TCC TTT GGG AGT GGG CCT GCG GTG  2388

E  W  I  P  K  Y  A  G  V  S  P  K  D  R  C  K  L  I  C  Q   663
GAA TGG ATT CCC AAG TAC GCT GGC GTC TCA CCA AAG GAC AGG TGC AAG CTC ATC TGC CAA  2448

A  K  G  I  G  Y  F  F  V  L  Q  P  K  V  V  D  G  T  P  C   683
```

FIGURE 1 (2 of 4)

```
GCC AAA GGC ATT GGC TAC TTC TTC GTT TTG CAG CCC AAG GTT GTA GAT GGT ACT CCA TGT  2508

 S  P  D  S  T  S  V  C  V  Q  G  Q  C  V  K  A  G  C  D  R   703
AGC CCA GAT TCC ACC TCT GTC TGT GTG CAA GGA CAG TGT GTA AAA GCT GGT TGT GAT CGC  2568

 I  I  D  S  K  K  K  F  D  K  C  G  V  C  G  G  N  G  S  T   723
ATC ATA GAC TCC AAA AAG AAG TTT GAT AAA TGT GGT GTT TGC GGG GGA AAT GGA TCT ACT  2628

 C  K  R  I  S  G  S  V  T  S  A  K  P  G  Y  H  D  I  I  T   743
TGT AAA AAA ATA TCA GGA TCA GTT ACT AGT GCA AAA CCT GGA TAT CAT GAT ATC ATC ACA  2688

 I  P  T  G  A  T  N  I  E  V  K  Q  R  N  Q  R  G  S  R  N   763
ATT CCA ACT GGA GCC ACC AAC ATC GAA GTG AAA CAG CGG AAC CAG AGG GGA TCC AGG AAC  2748

 N  G  S  F  L  A  I  K  A  A  D  G  T  Y  I  L  N  G  D  Y   783
AAT GGC AGC TTT CTT GCC ATC AAA GCT GCT GAT GGC ACA TAT ATT CTT AAT GGT GAC TAC  2808

 T  L  S  T  L  E  Q  D  I  M  Y  K  G  V  V  L  R  Y  S  G   803
ACT TTG TCC ACC TTA GAG CAA GAC ATT ATG TAC AAA GGT GTT GTC TTG AGG TAC AGC GGC  2868

 S  S  A  A  L  E  R  I  R  S  F  S  P  L  K  E  P  L  T  I   823
TCC TCT GCG GCA TTG GAA AGA ATT CGC AGC TTT AGC CCT CTC AAA GAG CCC TTG ACC ATC  2928

 Q  V  L  T  V  G  N  A  L  R  P  K  I  K  Y  T  Y  F  V  K   843
CAG GTT CTT ACT GTG GGC AAT GCC CTT CGA CCT AAA ATT AAA TAC ACC TAC TTC GTA AAG  2988

 K  K  K  E  S  F  N  A  I  P  T  F  S  A  W  V  I  E  E  W   863
AAG AAG AAG GAA TCT TTC AAT GCT ATC CCC ACT TTT TCA GCA TGG GTC ATT GAA GAG TGG  3048

 G  E  C  S  K  S  C  E  L  G  W  Q  R  R  L  V  E  C  R  D   883
GGC GAA TGT TCT AAG TCA TGT GAA TTG GGT TGG CAG AGA AGA CTG GTA GAA TGC CGA GAC  3108

 I  N  G  Q  P  A  S  E  C  A  K  E  V  K  P  A  S  T  R  P   903
ATT AAT GGA CAG CCT GCT TCC GAG TGT GCA AAG GAA GTG AAG CCA GCC AGC ACC AGA CCT  3168

 C  A  D  H  P  C  P  Q  W  Q  L  G  E  W  S  S  C  S  K  C   923
TGT GCA GAC CAT CCC TGC CCC CAG TGG CAG CTG GGG GAG TGG TCA TCA TGT TCT AAG ACC  3228

 C  G  K  G  Y  K  K  R  S  L  K  C  L  S  H  D  G  G  V  L   943
TGT GGG AAG GGT TAC AAA AAA AGA AGC TTG AAG TGT CTG TCC CAT GAT GGA GGG GTG TTA  3288

 S  H  E  S  C  D  P  L  K  K  P  K  H  F  I  D  F  C  T  M   963
TCT CAT GAG AGC TGT GAT CCT TTA AAG AAA CCT AAA CAT TTC ATA GAC TTT TGC ACA ATG  3348

 A  E  C  S  *
GCA GAA TGC AGT TAA                                                       968
                                                                         3363

GTGGTTTAAGTGGTGTTAGCTTTGAGGGCAAGGCAAAGTGAGGAAGGGCTGGTGCAGGGAAAGCAAGAAGGCTGGAGGG  3442

ATCCAGCGTATCTTGCCAGTAACCAGTGAGGTGTATCAGTAAGGTGGGATTATGGGGGTAGATAGAAAAGGAGTTGAAT  3521

CATCAGAGTAAACTGCCAGTTGCAAATTTGATAGGATAGTTAGTGAGGATTATTAACCTCTGAGCAGTGATATAGCATA  3600

ATAAAGCCCCGGGCATTATTATTATTATTTCTTTTGTTACATCTATTACAAGTTTAGAAAAAACAAAGCAATTGTCAAA  3679

AAAAGTTAGAACTATTACAACCCCTGTTTCCTGGTACTTATCAAATACTTAGTATCATGGGGGTTGGGAAATGAAAAGT  3758

AGGAGAAAAGTGAGATTTTACTAAGACCTGTTTTACTTTACCTCACTAACAATGGGGGGAGAAAGGAGTACAAATAGGA  3837
```

FIGURE 1 (3 OF 4)

TCTTTGACCAGCACTGTTTATGGCTGCTGTGGTTTCAGAGAATGTTTATACATTATTTCTACCGAGAATTAAAACTTCA 3916

GATTGTTCAACATGAGAGAAAGGCTCAGCAACGTGAAATAACGCAAATGGCTTCCTCTTTCCTTTTTTGGACCATCTCA 3995

GTCTTTATTTGTGTAATTCATTTTGAGGAAAAAACAACTCCATGTATTTATTCAAGTGCATTAAAGTCTACAATGGAAA 4074

AAAAGCAGTGAAGCATTACATGCTGGTAAAAGCTAGAGGAGACACAATGAGCTTAGTACCTCCAACTTCCTTTCTTTCC 4153

TACCATGTAACCCTGCTTTCGGAATATGGATGTAAAGAAGTAACTTGTGTCTCATGAAAATCAGTACAATCACACAAGG 4232

AGGATGAAACGCCCGAACAAAAATGAGGTGTGTAGAACAGGGTCCCACAGGTTTGGGGACATTGAGATCACTTGTCTTG 4311

TGGTGGGGAGGCTGCTGAGGGGTAGCAGGTCCATCTCCAGCAGCTGGTCCAACAGTCGTATCCTGGTGAATGTCTGTTC 4390

AGCTCTTCTGTGAGAATATGATTTTTTCCATATGTATATAGTAAAATATGTTACTATAAATTACATGTACTTTATAAGT 4469

ATTGGTTTGGGTGTTCCTTCCAAGAAGGACTATAGTTAGTAATAAATGCCTATAATAACATATTTATTTTTATACATTT 4548

ATTTCTAATGAAAAAAACTTTTAAATTATATCGCTTTTGTGGAAGTGCATATAAAATAGAGTATTTATACAATATATGT 4627

TACTAGAAATAAAAGAACACTTTTGGAAAAAAAAAAAAAAAAAAAAAAA 4676

FIGURE 1 (4 OF 4)

```
GTCGACCCACGCGTCCGGGGAGCAACGCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCGGGCGCGGACCCCAACCC  79

CGACCCAGAGCTTCTCCAGCGGCGGGGCAGCGAGCAGGGCTCCCCGCCTTAACTTCCTCGCGGGGCCCAGCCACCTTC  158

GGGAGTCCGGGTTGCCCACCTGCAAACTCTCCGCCTTCTGCACCTGCCACCCCTGAGCCAGCGCGGGCGCCCGAGCGAG  237

   M  A  N  A  G  L  Q  L  L  G  F  I  L  A  F  L  G  W  I    19
TC ATG GCC AAC GCG GGG CTG CAG CTG TTG GGC TTC ATT CTC GCC TTC CTG GGA TGG ATC  296

 G  A  I  V  S  T  A  L  P  Q  W  R  I  Y  S  Y  A  G  D  N    39
GGC GCC ATC GTC AGC ACT GCC CTG CCC CAG TGG AGG ATT TAC TCC TAT GCC GGC GAC AAC  356

 I  V  T  A  Q  A  M  Y  E  G  L  W  M  S  C  V  S  Q  S  T    59
ATC GTG ACC GCC CAG GCC ATG TAC GAG GGG CTG TGG ATG TCC TGC GTG TCG CAG AGC ACC  416

 G  Q  I  Q  C  K  V  F  D  S  L  L  N  L  S  S  T  L  Q  A    79
GGG CAG ATC CAG TGC AAA GTC TTT GAC TCC TTG CTG AAT CTG AGC AGC ACA TTG CAA GCA  476

 T  R  A  L  M  V  V  G  I  L  L  G  V  I  A  I  F  V  A  T    99
ACC CGT GCC TTG ATG GTG GTT GGC ATC CTC CTG GGA GTG ATA GCA ATC TTT GTG GCC ACC  536

 V  G  M  K  C  M  K  C  L  E  D  D  E  V  Q  K  M  R  M  A   119
GTT GGC ATG AAG TGT ATG AAG TGC TTG GAA GAC GAT GAG GTG CAG AAG ATG AGG ATG GCT  596

 V  I  G  G  A  I  F  L  L  A  G  L  A  I  L  V  A  T  A  W   139
GTC ATT GGG GGT GCG ATA TTT CTT CTT GCA GGT CTG GCT ATT TTA GTT GCC ACA GCA TGG  656

 Y  G  N  R  I  V  Q  E  F  Y  D  P  M  T  P  V  N  A  R  Y   159
TAT GGC AAT AGA ATC GTT CAA GAA TTC TAT GAC CCT ATG ACC CCA GTC AAT GCC AGG TAC  716

 E  F  G  Q  A  L  F  T  G  W  A  A  A  S  L  C  L  L  G  G   179
GAA TTT GGT CAG GCT CTC TTC ACT GGC TGG GCT GCT GCT TCT CTC TGC CTT CTG GGA GGT  776

 A  L  L  C  C  S  C  P  R  K  T  T  S  Y  P  T  P  R  P  Y   199
GCC CTA CTT TGC TGT TCC TGT CCC CGA AAA ACA ACC TCT TAC CCA ACA CCA AGG CCC TAT  836

 P  K  P  A  P  S  S  G  K  D  Y  V  *                       212
CCA AAA CCT GCA CCT TCC AGC GGG AAA GAC TAC GTG TGA                875

CACAGAGGCAAAAGGAGAAAATCATGTTGAAACAAACCGAAAATGGACATTGAGATACTATCATTAACATTAGGACCTT  954

AGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAACAAAAAACCCATGTGTTAAAATACTC  1033

AGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTCCTCAATATAGGAGGGAAGATTTTTCCATTTGTATTACTGCT  1112

TCCCATTGAGTAATCATACTCAACTGGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTC  1191

TATTAAAAATAGACAGTAAAATACTATTCTCATTATGTTGATACTAGCATACTTAAAATATCTCTAAAATAGGTAAATG  1270

TATTTAATTCCATATTGATGAAGATGTTTATTGGTATATTTTCTTTTTCGTCTATATATACATATGTAACAGTCAAATA  1349

TCATTTACTCTTCTTCATTAGCTTTGGGTGCCTTTGCCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTTCAAT  1428

TCTTCATGCGTGCCCTTTTCATATACTTATTTTATTTTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTT  1507

ATTTGTTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTCATAGCCTACATTTTAGTTTCTAAAGCCAAG  1586

AAGAATTTATTACAAATCAGAACTTTGGAGGCAAATCTTTCTGCATGACCAAAGTGATAAATTCCTGTTGACCTTCCCA  1665
```

Fig. 2 (1 of 2)

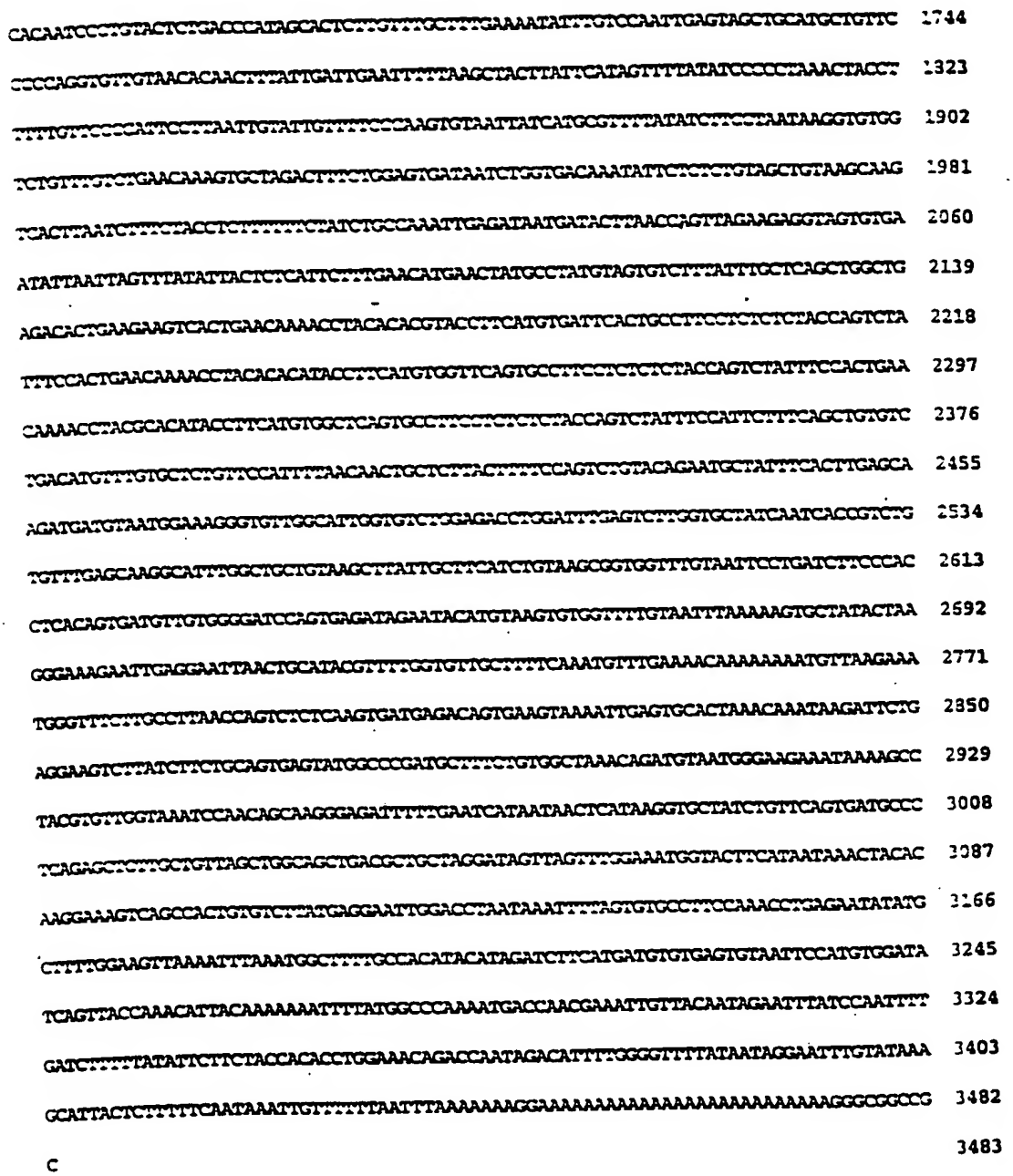

CACAATCCCTGTACTCTGACCCATAGCACTCTTGTTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTC 1744
CCCCAGGTGTTGTAACACAACTTTATTGATTGAATTTTAAGCTACTTATTCATAGTTTTATATCCCCCTAAACTACCT 1323
TTTGTTCCCCCATTCCTTAATTGTATTGTTTTCCCAAGTGTAATTATCATGCGTTTTATATCTTCCTAATAAGGTGTGG 1902
TCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATAATCTGGTGACAAATATTCTCTCTGTAGCTGTAAGCAAG 1981
TCACTTAATCTTTCTACCTCTTTTTTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGA 2060
ATATTAATTAGTTTATATTACTCTCATTCTTTGAACATGAACTATGCCTATGTAGTGTCTTTATTTGCTCAGCTGGCTG 2139
AGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATTCACTGCCTTCCTCTCTCTACCAGTCTA 2218
TTTCCACTGAACAAAACCTACACACATACCTTCATGTGGTTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCACTGAA 2297
CAAAACCTACGCACATACCTTCATGTGGCTCAGTGCCTTCCTCTCTCTACCAGTCTATTTCCATTCTTTCAGCTGTGTC 2376
TGACATGTTTGTGCTCTGTTCCATTTTAACAACTGCTCTTACTTTTCCAGTCTGTACAGAATGCTATTTCACTTGAGCA 2455
AGATGATGTAATGGAAAGGGTGTTGGCATTGGTGTCTGGAGACCTGGATTTGAGTCTTGGTGCTATCAATCACCGTCTG 2534
TGTTTGAGCAAGGCATTTGGCTGCTGTAAGCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAATTCCTGATCTTCCCAC 2613
CTCACAGTGATGTTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAATTTAAAAAGTGCTATACTAA 2692
GGGAAAGAATTGAGGAATTAACTGCATACGTTTTGGTGTTGCTTTTCAAATGTTTGAAAACAAAAAAAAATGTTAAGAAA 2771
TGGGTTTCTTGCCTTAACCAGTCTCTCAAGTGATGAGACAGTGAAGTAAAATTGAGTGCACTAAACAAATAAGATTCTG 2850
AGGAAGTCTTATCTTCTGCAGTGAGTATGGCCCGATGCTTTCTGTGGCTAAACAGATGTAATGGGAAGAAATAAAAGCC 2929
TACGTGTTGGTAAATCCAACAGCAAGGGAGATTTTTGAATCATAATAACTCATAAGGTGCTATCTGTTCAGTGATGCCC 3008
TCAGAGCTCTTGCTGTTAGCTGGCAGCTGACGCTGCTAGGATAGTTAGTTTGGAAATGGTACTTCATAATAAACTACAC 3087
AAGGAAAGTCAGCCACTGTGTCTTATGAGGAATTGGACCTAATAAATTTTAGTGTGCCTTCCAAACCTGAGAATATATG 3166
CTTTTGGAAGTTAAAATTTAAATGGCTTTTGCCACATACATAGATCTTCATGATGTGTGAGTGTAATTCCATGTGGATA 3245
TCAGTTACCAAACATTACAAAAAAATTTTATGGCCCAAAATGACCAACGAAATTGTTACAATAGAATTTATCCAATTTT 3324
GATCTTTTTATATTCTTCTACCACACCTGGAAACAGACCAATAGACATTTTGGGGTTTTATAATAGGAATTTGTATAAA 3403
GCATTACTCTTTTTCAATAAATTGTTTTTTAATTTAAAAAAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCG 3482
C 3483

Fig. 2 (2 of 2)

```
GTCGACCCACGCGTCCGGCTGCGAGAACCTTTGCACGGGCACAAACTACGGGGACGATTTCTGATTGATTTTTGGCGCT   79

                                M  G  L  W  G  Q  S  V  P  T  A  S  S      13
TTCGATCCACCCTCGTCCCTTCTC ATG GGA CTT TGG GGA CAA AGC GTC CCG ACC GCC TCG AGC      142

 A  R  A  G  R  Y  P  G  A  R  T  A  S  G  T  R  P  W  L  L      33
GCT CGA GCA GGG CGC TAT CCA GGA GCC AGG ACA GCG TCG GGA ACC AGA CCA TGG CTC CTG   202

 D  E  K  I  L  K  F  V  V  F  I  V  A  V  L  L  P  V  R  V      53
GAC TCC AAG ATC CTT AAG TTC GTC GTC TTC ATC GTC GCG GTT CTG CTG CCG GTC CGG GTT   262

 D  S  A  T  I  P  R  Q  D  E  V  P  Q  Q  T  V  A  P  Q  Q      73
GAC TCT GCC ACC ATC CCC CGG CAG GAC GAA GTT CCC CAG CAG ACA GTG GCC CCA CAG CAA   322

 Q  R  R  S  L  K  E  E  E  C  P  A  G  S  H  R  S  E  Y  T      93
CAG AGG CGC AGC CTC AAG GAG GAG GAG TGT CCA GCA GGA TCT CAT AGA TCA GAA TAT ACT   382

 G  A  C  N  P  C  T  E  G  V  D  Y  T  I  A  S  N  N  L  P     113
GGA GCC TGT AAC CCG TGC ACA GAG GGT GTG GAT TAC ACC ATT GCT TCC AAC AAT TTG CCT   442

 S  C  L  L  C  T  V  C  K  S  G  Q  T  N  K  S  S  C  T  T     133
TCT TGC CTG CTA TGT ACA GTT TGT AAA TCA GGT CAA ACA AAT AAA AGT TCC TGT ACC ACG   502

 T  R  D  T  V  C  Q  C  E  K  G  S  F  Q  D  K  N  S  P  E     153
ACC AGA GAC ACC GTG TGT CAG TGT GAA AAA GGA AGC TTC CAG GAT AAA AAC TCC CCT GAG   562

 M  C  R  T  C  R  T  G  C  P  R  G  M  V  K  V  S  N  C  T     173
ATG TGC CGG ACG TGT AGA ACA GGG TGT CCC AGA GGG ATG GTC AAG GTC AGT AAT TGT ACG   622

 P  R  S  D  I  K  C  K  N  E  S  A  A  S  S  T  G  K  T  P     193
CCC CGG AGT GAC ATC AAG TGC AAA AAT GAA TCA GCT GCC AGT TCC ACT GGG AAA ACC CCA   682

 A  A  E  E  T  V  T  T  I  L  G  M  L  A  S  P  Y  H  Y  L     213
GCA GCG GAG GAG ACA GTG ACC ACC ATC CTG GGG ATG CTT GCC TCT CCC TAT CAC TAC CTT   742

 I  I  I  V  V  L  V  I  I  L  A  V  V  V  V  G  F  S  C  R     233
ATC ATC ATA GTG GTT TTA GTC ATC ATT TTA GCT GTG GTT GTG GTT GGC TTT TCA TGT CGG   802

 K  K  F  I  S  Y  L  K  G  I  C  S  G  G  G  G  G  P  E  R     253
AAG AAA TTC ATT TCT TAC CTC AAA GGC ATC TGC TCA GGT GGT GGA GGA GGT CCC GAA CGT   862

 V  H  R  V  L  F  R  R  R  S  C  P  S  R  V  P  G  A  E  D     273
GTG CAC AGA GTC CTT TTC CGG CGG CGT TCA TGT CCT TCA CGA GTT CCT GGG GCG GAG GAC   922

 N  A  R  N  E  T  L  S  N  R  Y  L  Q  P  T  Q  V  S  E  Q     293
AAT GCC CGC AAC GAG ACC CTG AGT AAC AGA TAC TTG CAG CCC ACC CAG GTC TCT GAG CAG   982

 E  I  Q  G  Q  E  L  A  E  L  T  G  V  T  V  E  S  P  E  E     313
GAA ATC CAA GGT CAG GAG CTG GCA GAG CTA ACA GGT GTG ACT GTA GAG TCG CCA GAG GAG  1042

 P  Q  R  L  L  E  Q  A  E  A  E  G  C  Q  R  R  R  L  L  V     333
CCA CAG CGT CTG CTG GAA CAG GCA GAA GCT GAA GGG TGT CAG AGG AGG AGG CTG CTG GTT  1102

 P  V  N  D  A  D  S  A  D  I  S  T  L  L  D  A  S  A  T  L     353
CCA GTG AAT GAC GCT GAC TCC GCT GAC ATC AGC ACC TTG CTG GAT GCC TCG GCA ACA CTG  1162

 E  E  G  H  A  K  E  T  I  Q  D  Q  L  V  G  S  E  K  L  F     373
GAA GAA GGA CAT GCA AAG GAA ACA ATT CAG GAC CAA CTG GTG GGC TCC GAA AAG CTC TTT  1222
```

FIG. 3 (sheet 1 of 2)

```
Y  E  E  D  E  A  G  S  A     S  C  L  *                                  387
TAT GAA GAA GAT GAA GCA GGC TCT GCT ACG TCC TGC CTG TGA                    1264

AAGAATCTGTTCAGGAAACCAGAGCTTCCCTCATTTACCTTTTCTCCTACAAAGGGAAGCAGCCTGGAGAAACAGTCC  1343

AGTACTTGACCCATGCCCCAACAAACTCTACTATCCAATATGGGGCAGCTTACCAATGGTCCTAGAACTTTGTTAACGC  1422

ACTTGGAGTAATTTTTATGAAATACTGCGTGTGATAAGCAAACGGGAGAAATTTATATCAGATTCTTGGCTGCATAGTT  1501

ATACGATTGTGTATTAAGGGTCGTTTTAGGCCACATGCGGTGGCTCATGCCTGTAATCCCAGCACTTTGATAGGCTGAG  1580

CCAGGTGGATTGCTTGAGCTCGGGAGTTTGAGACCAGCCTGATCAACACAGTGAAACTCCATCTCAATTTAAAAAGAAA  1659

AAAAGTGGTTTTAGGATGTCATTCTTTGCAGTTCTTCATCATGAGACAAGTCTTTTTTTCTGCTTCTTATATTGCAAGC  1738

TCCATCTCTACTGGTGTGTGCATTTAATGACATCTAACTACAGATGCCGCACAGCCACAATGCTTTGCCTCTATAATTTT  1817

TTAACTTTAGAACGGGATTATCTTGTTATTACCTGTATTTTCAGTTTCGGATATTTTTGACTTAATGATGAGATTATCA  1896

AGACGTAGCCCTATGCTAAGTCATGAGCATATGGACTTACGAGGGTTCGACTTAGAGTTTTGAGCTTTAAGATACGATT  1975

ATTGGGGCTTACCCCCACCTTAATTAGAGAAACATTTATATTGCTTACTACTGTAGGCTGTACATCTCTTTTCCGATTT  2054

TTGTATAAATGATGTAAACATGGAAAAACTTTAGGAAATGCACTTATTAGGCTGTTTACATGGGTTGCCTGGATACAAAT  2133

CAGCAGTCAAAAATGACTAAAAATATAACTAGTGACGGAGGGAGAAATCCTCCCTCTGTGGGAGGCACTTACTGCATTC  2212

CAGTTCTCCCTTCCTGCGCCCTGAGACTGGACCAGGGTTTGATGGCTGGCAGCTTCTCAAGGGGCAGCTTGTCTTACTTG  2291

TTAATTTTAGAGGTATATAGCCATATTTATTTATAAATAAATATTTATTTATTTATTTATAAGTAGATGTTTACATATG  2370

CCCAGGATTTTGAAGAGCCTGGTATCTTTGGGAAGCCATGTGTCTGGTTTGTCGTGCTGGGACAGTCATGGGACTGCAT  2449

CTTCCGACTTGTCCACAGCAGATGAGGACAGTGAGAATTAAGTTAGATCCGAGACTGCGAAGAGCTTCTCTTTCAAGCG  2528

CCATTACAGTTGAACGTTAGTGAATCTTGAGCCTCATTTGGGCTCAGGGCAGAGCAGGTGTTTATCTGCCCCGGCATCT  2607

GCCATGGCATCAAGAGGGAAGAGTGGACGGTGCTTGGGAATGGTGTGAAATGGTTGCCGACTCAGGCATGGATGGGCCC  2686

CTCTCGCTTCTGGTGGTCTGTGAACTGAGTCCCTGGGATGCCTTTTAGGGCAGAGATTCCTGAGCTGCGTTTTAGGGTA  2765

CAGATTCCCTGTTTGAGGAGCTTGGCCCCTCTGTAAGCATCTGACTCATCTCAGAGATATCAATTCTTAAACACTGTGA  2844

CAACAGGATCTAAAATGGCTGACACATTTGTCCTTGTGTCACGTTCCATTATTTTATTTAAAAACCTCAGTAATCGTTT  2923

TAGCTTCTTTCCAGCAAACTCTTCTCCACAGTAGCCCAGTCGTGGTAGGATAAATTACGGATATAGTCATTCTAGGGGT  3002

TTCAGTCTTTTCCATCTCAAGGCATTGTGTGTTTTGTTCCGGGACTGGTTTGGCTGGGACAAAGTTAGAACTGCCTGAA  3081

GTTCGCACATTCAGATTGTTGTGTCCATGGAGTTTTAGGAGGGGATGGCCTTTCCGGTCTTCGCACTTCCATCCTCTCC  3160

CACTTCCATCTGGCGTCCCACACCTTGTCCCCTGCACTTCTGGATGACACAGGGTGCTGCTGCCTCCTAGTCTTTGCCT  3239

TTGCTGGGCCTTCTGTGCAGGAGACTTGGTCTCAAAGCTCAGAGAGAGCCAGTCCGGTCCCAGCTCCTTTGTCCCTTCC  3318

TCAGAGGCCTTCCTTGAAGATGCATCTAGACTACCAGCCTTATCAGTGTTTAAGCTTATTCCTTTAACATAAGCTTCCT  3397

GACAACATGAAATTGTTGGGGTTTTTTGGCGTTGGTTGAATTGTTTAGGTTTTGCTTTATACCCGGGCCAAATAGCACA  3476

TAACACCTGGTTATATATGAAATACTCATATGTTTATGACCAAAATAAATATGAAACCTCATATTAAAAAAAAAAAAAA  3555

AAAAGGGCGGCCGC                                                             3569
```

FIG. 3 (sheet 2 of 2)

GCCTTTTAGGGCAGAGATTCCTGAGCTGCGTTTTAGGGTACAGATTCCCT
GTTTGAGGAGCTTGGCCCCTCTGTAAGCATCTGACTCATCTCAGAGATAT
CAATTCTTAAACACTGTGACAACAGGATCTAAAATGGCTGACACATTTGT
CCTTGTGTCACGTTCCATTATTTTATTTAAAAACCTCAGTAATCGTTTTA
GCTTCTTTCCAGCAAACTCTTCTCCACAGTAGCCCAGTCGTGGTAGGATA
AATTACGGATATAGTCATTCTAGGGGTTTCAGTCTTTTCCATCTCAAGGC
ATTGTGTGTTTTGTTCCGGGACTGGTTTGGCTGGGACAAAGTTAGAACTG
CCTGAAGTTCGCACATTCAGATTGTTGTGTCCATGGAGTTTTAGGAGGGG
ATGGCCTTTCCGGTCTTCGCACTTCCATCCTCTCCCCACTTCCATCTGGCG
TCCCACACCTTGTCCCCTGCACTTCTGGATGACACAGGGTGCTGCTGCCT
CCTAGTCTTTGCCTTTGCTGGGCCTTCTGTGCAGGAGACTTGGTCTCAAA
GCTCAGAGAGAGCCAGTCCGGTCCCAGCTCCTTTGTCCCTTCCTCAGAGG
CCTTCCTTGAAGATGCATCTAGACTACCAGCCTTATCAGTGTTTAAGCTT
ATTCCTTTAACATAAGCTTCCTGACAACATGAAATTGTTGGGGTTTTTTG
GCGTTGGTTGATTTGTTTAGGTTTTGCTTTATACCCGGGCCAAATAGCAC
ATAACACCTGGTTATATATGAAATACTCATATGTTTATGACCAAAATAAA
TATGAAACCTCATATTAAAAAAAAAAAAAAAAAAAGGGCGGCCG

Fig. 4

56.2% identity in 176 aa overlap

```
          30        40        50        60        70        80
TRAIL AAATPSKVWGSSAGRIEPRGGGRGALPTSMGQHGPSA-RARAGRAPGPRPAREASPRLRV
                                    ::  :.:  ::::: ::.: :  . :.:
T74                                 MGLWGQSVPTASSARAGRYPGARTASGTRPWLLD
                                              10        20        30

          90        100       110       120       130       140
TRAIL HKTFKFVV--VGVLLQVVPSSAATIK---LHDQSIGTQQWEHSPLGELCPPGSHRSERPG
       :..:::  :.:::.:  .::.. .   . .:... ::  .: X: ::.::::::  :
T74    SKILKFVVFIVAVLLPVRVDSATIPRQDEVPQQTVAPQQQRRSLKEEECPAGSHRSEYTG
          40        50        60        70        80        90

              150       160       170       180       190       200
TRAIL ACNRCTEGVGYTNASNNLFACLPCTACKSDEEERSPCTTTRNTACQCKPGTFRNDNSAEM
      ::: :::::.:: ::::: .:: ::.:::.. ..:.::::::.:.::: :.:.. ::.::
T74   ACNPCTEGVDYTIASNNLPSCLLCTVCKSGQTNKSSCTTTRDTVCQCEKGSFQDKNSPEM
          100       110       120       130       140       150

              210       220       230       240       250       260
TRAIL CRKCSTGCPRGMVKVKDCTPWSDIECVHKESGNGHNIWVILVVTLVVPLLLVAVLIVCCC
      :: : :::::::::: .::::.::X :
T74   CRTCRTGCPRGMVKVSNCTPRSDIKCKNESAASSTGKTPAAEETVTTILGMLASPYHYLI
          160       170       180       190       200       210
```

Fig. 5

```
   V  R  N  K  T  L  P  W  S  P  C  S  A  V  Y  L  T  E  L      19
GC GTC CGG AAC AAG ACG CTG CCC TGG TCT CCC TGC AGT GCT GTC TAC CTC ACG GAG CTC      59

L  D  D  G  H  G  D  C  L  L  D  D  G  H  S  T  L  Y  E  L      39
CTG GAT GAT GGT CAC GGA GAC TGC CTC CTG GAT GAT GGC CAC AGC ACC CTC TAT GAG CTG     119

D  Q  Q  C  K  Q  I  F  G  P  D  F  R  H  C  P  N  T  S  V      59
GAC CAG CAG TGC AAG CAG ATC TTT GGG CCT GAT TTC CGA CAC TGC CCC AAC ACC TCT GTG     179

E  D  I  C  V  Q  L  W  C  R  H  R  D  S  D  E  P  I  C  H      79
GAG GAC ATC TGT GTC CAG CTC TGG TGC CGT CAT CGG GAT AGT GAT GAG CCC ATT TGC CAC     239

T  K  N  G  S  L  L  W  A  D  G  T  P  C  G  P  G  H  L  C      99
ACA AAG AAT GGC AGC TTG CTC TGG GCT GAT GGT ACG CCC TGT GGC CCT GGG CAC CTG TGC     299

L  D  G  S  C  V  L  R  E  E  V  E  N  P  K  A  V  V  D  G     119
CTG GAT GGT AGC TGT GTG CTC CGG GAG GAA GTA GAG AAT CCC AAG GCT GTG GTA GAT GGA     359

D  W  G  P  W  G  P  W  G  Q  C  S  R  T  C  G  G  G  I  Q     139
GAC TGG GGT CCC TGG GGA CCC TGG GGA CAA TGT TCT CGC ACC TGT GGT GGA GGG ATA CAG     419

F  S  N  R  E  C  D  N  P  A  P  Q  N  G  G  R  F  C  L  G     159
TTT TCG AAC CGT GAG TGT GAT AAT CCA GCA CCT CAG AAT GGA GGA AGA TTT TGC CTG GGA     479

E  R  V  K  Y  Q  S  C  K  T  E  E  C  P  P  N  G  K  S  F     179
GAG AGA GTC AAG TAC CAA TCT TGC AAG ACA GAG GAA TGT CCA CCA AAT GGA AAA AGC TTC     539

R  E  Q  Q  C  E  K  Y  N  A  Y  N  H  T  D  L  D  G  N  F     199
AGG GAG CAG CAG TGT GAA AAA TAT AAT GCC TAC AAC CAC ACG GAC CTG GAT GGG AAT TTC     599

L  Q  W  V  P  K  Y  S  G  V  S  P  R  D  R  C  K  L  F  C     219
CTT CAG TGG GTC CCC AAA TAC TCA GGA GTG TCC CCC CGA GAC CGA TGC AAA CTG TTT TGC     659

R  A  R  G  R  S  E  F  K  V  F  E  T  K  V  I  D  G  T  L     239
AGA GCC CGT GGG AGG AGT GAG TTC AAA GTG TTT GAA ACT AAG GTG ATC GAT GGC ACT CTG     719

C  G  P  D  T  L  A  I  C  V  R  G  Q  C  V  K  A  G  C  D     259
TGC GGA CCG GAT ACT CTG GCC ATC TGT GTG CGG GGA CAG TGC GTT AAG GCT GGC TGT GAC     779

H  V  V  N  S  P  K  K  L  D  K  C  G  V  C  G  G  K  G  T     279
CAT GTG GTG AAC TCA CCT AAG AAG CTG GAC AAG TGT GGG GTG TGT GGG GGC AAA GGC ACT     839

A  C  R  K  V  S  G  S  F  T  P  F  S  Y  G  Y  N  D  I  V     299
GCC TGT AGG AAG GTC TCA GGT TCT TTC ACC CCC TTC AGT TAT GGC TAC AAT GAC ATT GTC     899

T  I  P  A  G  A  T  N  I  D  V  K  Q  R  S  H  P  G  V  Q     319
ACC ATC CCA GCT GGT GCC ACA AAT ATT GAT GTG AAA CAA CGG AGC CAC CCA GGG GTC CAG     959

N  D  G  S  Y  L  A  L  K  T  A  N  G  Q  Y  L  L  N  G  N     339
AAT GAC GGC AGC TAC CTG GCA CTG AAG ACA GCC AAT GGG CAG TAC CTG CTC AAT GGT AAC    1019

L  A  I  S  A  I  E  Q  D  I  L  M  K  G  T  I  L  K  Y  S     359
CTA GCC ATC TCT GCC ATA GAG CAA GAC ATC TTG ATG AAG GGG ACC ATC CTA AAG TAC AGT    1079

G  S  M  A  T  L  E  R  L  Q  S  F  Q  A  L  P  E  P  L  T     379
GGT TCC ATG GCC ACC CTG GAG CGG CTG CAG AGC TTC CAA GCC CTC CCT GAG CCT CTT ACA    1139
```

Fig. 6 (1 of 2)

```
V  Q  L  L  T  V  S  G  E  V  F  P  P  K  V  K  Y  T  F  F   399
GTA CAG CTC CTG ACT GTG TCT GGT GAG GTC TTC CCT CCA AAA GTC AAA TAT ACC TTC TTC  1199

V  P  N  D  T  D  F  N  V  Q  S  S  K  E  R  A  S  T  N  I   419
GTC CCC AAT GAC ACG GAC TTC AAC GTG CAG AGT AGC AAA GAA AGA GCA AGC ACC AAC ATC  1259

I  Q  S  L  P  Y  A  E  W  V  L  G  D  W  S  E  C  P  S  T   439
ATT CAG TCC TTG CCC TAT GCA GAG TGG GTG CTG GGG GAC TGG TCT GAA TGT CCA AGC ACA  1319

C  G  G  G  W  Q  R  R  T  V  E  C  R  D  P  S  G  Q  A  S   459
TGT GGA GGT GGC TGG CAG CGG CGG ACT GTG GAA TGC AGG GAC CCC TCA GGT CAG GCC TCT  1379

D  T  C  D  E  A  L  K  P-  E  D  A  K  P  C  G  S  Q  P  C   479
GAC ACC TGT GAT GAG GCT CTG AAA CCT GAG GAT GCC AAG CCC TGT GGA AGC CAG CCA TGT  1439

L  L  *                                                        482
CTC CTC TGA                                                    1448

TCCCCTTGGTGGACATGTCTAAGGCTTATGGATTTGGGCTACTGGCGTACAGACAAAGGTCTCCTCTGAGGTGACACTA  1527

CATATCAAGATGGCATGGCCCTTCCAGGCCTTCTATTACTACAACCCTTTGGGTACCACCTAATTCATAAGGAAGAGAG  1606

AAGAGGATGTAAGGGTAACAGACTGTAAAGTTGACTGTCTAGTGGACTGGACCTTGTTTATGACCAAGAAGATGGGATA  1685

GGTTAAAAGGTAAAAGTGTGCTTATTGATCCAAAGGTGAGATTTCAGAACCAGCCTCTTTGCAAAGGACTAGAAAGGTT  1764

AAATGAGAAAGAAGAATTTTTTTTCTCTTTGGTTTCTCCAATAATCAATCTACCTCACAGCGGGAGGAACTTGGTGTAT  1843

AAGGCCAGGTGTTAGTGGTGAGTGCCAAGGCACTCTCCATAGATATCTTCGAGCCATCTTCAGAAATGGCCATGGCTGT  1922

TTTCAGTATTAAAACTCTGTTGTCTCAAAAGGTGGTGGTGTCCATCACAGGGTTATAGAAAGCCACTTGTTCTCAGGCT  2001

GCCTCCTGCTGGGGCGGACCCCTTTCAAGTATTTATGCAAATATGTTTCTGAACTAAAGTGTGATCTTACACCAAAAAA  2080

AAAAAAAAAAAAAAAAAAAAAAAAAGGCGGCCGC                              2114
```

Fig. 6 (2 of 2)

>lthxe29c9F. 1448 bases, 3233 checksum.

GTCGACCCACGCGTCCGGGGGGAAGCTTGCCAGCAGATCTGCAGCTGCCAA
AATGGGGCAGACTGTGACAGTGTGACTGGAAAGTGCACCTGTGCCCCAGG
ATTCAAAGGAATTGACTGCTCTACCCCATGCCCTCTGGGAACCTATGGGA
TAAACTGTTCCTCTCGCTGTGGCTGTAAAAATGATGCAGTCTGCTCTCCT
GTGGACGGGTCTTGTACTTGCAAGGCAGGCTGGCACGGGGTGGACTGCTC
CATCAGATGTCCCAGTGGCACATGGGGCTTTGGCTGTAACTTAACATGCC
AGTGCCTCAACGGGGGAGCCTGCAACACCCTGGACGGGACCTGCACGTGT
GCACCTGGATGGCGCGGGGAGAAATGCGAACTTCCCTGCCAGGATGGCAC
GTACGGGCTGAACTGTGCTGAGCGCTGCGACTGCAGCCACGCAGATGGCT
GCCACCCTACCACGGGCCATTGCCGCTGCCTCCCCGGATGGTCAGGTGTC
CACTGTGACAGCGTGTGTGCTGAGGGACGCTGGGGCCCCAACTGCTCCCT
GCCCTGCTACTGTAAAAATGGGGCTTCATGCTCCCCTGATGATGGCATCT
GCGAGTGTGCACCAGGCTTCCGAGGCACCACTTGTCAGAGGATCTGCTCC
CCTGGTTTTTATGGGCATCGCTGCAGCCAGACATGCCCACAGTGCGTTCA
CAGCAGCGGGCCCTGCCACCACATCACCGGCCTGTGTGACTGCTTGCCTG
GCTTCACAGGCGCCCTCTGCAATGAAGTGTGTCCCAGTGGCAGATTTGGG
AAAAACTGTGCAGGAATTTGTACCTGCACCAACAACGGAACCTGTAACCC
CATTGACAGATCTTGTCAGTGTTACCCCGGTTGGATTGGCAGTGACTGCT
CTCAACCATGTCCACCTGCCCACTGGGGCCCAAACTGCATCCACACGTGC
AACTGCCATAATGGAGCTTTCTGCAGCGCCTACGATGGGGAATGTAAATG
CACTCCTGGCTGGACAGGGCTCTACTGCACTCAGAGATGTCCTCTAGGGT
TTTATGGAAAAGATTGTGCACTGATATGCCAATGTCAAAACGGAGCTGAC
TGCGACCACATTTCTGGGCAGTGTACTTGCCGCACTGGATTCATGGGACG
GCACTGTGAGCAGAAGTGCCCTTCAGGAACATATGGCTATGGCTGTCGCC
AGATATGTGATTGTCTGAACAACTCCACCTGCGACCACATCACTGGGACC
TGTTACTGCAGCCCCGGATGGAAGGGAGCGAGATGTGATCAAGCTGGTGT
TATCATAGTTGGAAATCTGAACAGCTTAAGCCGAACCAGTACTGCTCTCC
CTGCTGATTCCTACCAAATCGGGGCCATTGCAGGCATCATCATTCTTGTC
CTAGTTGTTCTCTTCCTACTGGCATTGTTCATTATTTATAGACACAGC

Fig. 7

>fthxe29c9R, 1578 bases, 5059 checksum.

NAGCCCAACAGGAATGTTCTATGAAAGTGAACCTAACAGTGAGTGTTGTT
CCCAAGGAGTATTCAGCAATAATGGGCGTCTNTCCCAAGGATCCATATGA
CCTCCCAAAGAACAGTCACATCCCTTGTCATTATGACCTGCTGCCAGTCC
GAGACAGTTCATCCTCCCCTAAGCAAGAGGACAGTGGAGGTAGCAGCAGC
AACAGCAGCAGCAGCAGTGAATGACACCAAAGGACCGCTTGGTAGCCACT
GGAACCCTTTCCAGAACTGCTGTTTGGTTCTTCTCCATCCTCAATTTTGC
CACTTTCATGTGAATGTTAGTCAATTCGGTGGGCAATTTTTGGACATGAA
CCAGAAAGCTGAAAGCTGAGGCTGACACGGACTGTAGGTGCTTTTTGTTC
AGGTGGATTCGAAGGAGTTAGAGATGTGATTTGCCATTGCTGTTAGTTTT
AGAACTATACCCGTGAAGCATGACTTATTGTAAGATGTTGGCTGAAAGCA
TGAACTTGCAGAACTCCCTCGGAGACGCAGGTTGCAGTGGACATTGGGAT
TGTTGCTTGAAAAATTAAAATTTGAATATTTTCTCTCTCATTTGCATCAT
ACAGCTCTACCTAGGATTGTACAGTTTACCATAAAATTTACTTCATGAAA
GTGGGAATCACTGAACATGTAGAAGACAAGGAACATATTGTTAACTCCTG
ATTCTTAACTTTATTCAACTGGACTCAGAATTGTAGGGATAATATGAATG
CAGGAGGAAACATTCTGTCAGGCGGTATGACTGGACAGACTTTGAATATA
CTCTAAAAGTGGACAGAAAATTTACGAAAATCTTAGATTTTGTTTAGAAT
GAGAAAATATACAATTAGAATTATTTTAGAAATAGTAGGAAGTATTGCAG
AAGTCAATACACAAATGTGCCAGGCAGAGGTGGTTTTCTCTGTTTGACTC
TCAACCAACTTCAGATCTATGACATTATTCTGATCACTGGCTCCATCATA
CATATTCACCACTTGAGATTCATAACATATCAATAGTTATTTCATAAATA
TAGAAATGAAATAATTTTATTTTTGACAGACTGGATGGAATGAGTGTGTA
ATGATTGATAAAGGTTGTAAATTTTAAATGCAAGATGACGCTTACGTTCT
GTAAACCATTAGTAATACATGCTGTAATATAGAATTAGTGGAACATTTTG
ATTAATCTTTCCCTAGAAGTGACTGAAATATTTTTGTGCATATTTGAGAA
AGGGAACTTTCCTTTTATTAATTGTCAATTTAGAGAAACTATGCTTAAGC
TGGTCTTTTGCATTGCTAATGTGACATGTACCCAACTTTTCATTAATTTG
TATTTCCATTTTTAAATTGCATATTCTATGTTTTGTAGTGTTTGGATTGT
TAATGAAAAAATATTATATGTTCGTTATTCCTTGTATTATTGCCACTTAT
CTTTTGCTTGATAAAAATGCGTTGTTCTTTTTTCTTTTGGAGGGACAAGA
TGAAAATATATAATTTGAATTGATTAAAATTGGTCGTTACTAAAATAGTA
TAGTAAAAAAAAAAAAAAAGGGCGGCCG

Fig. 8

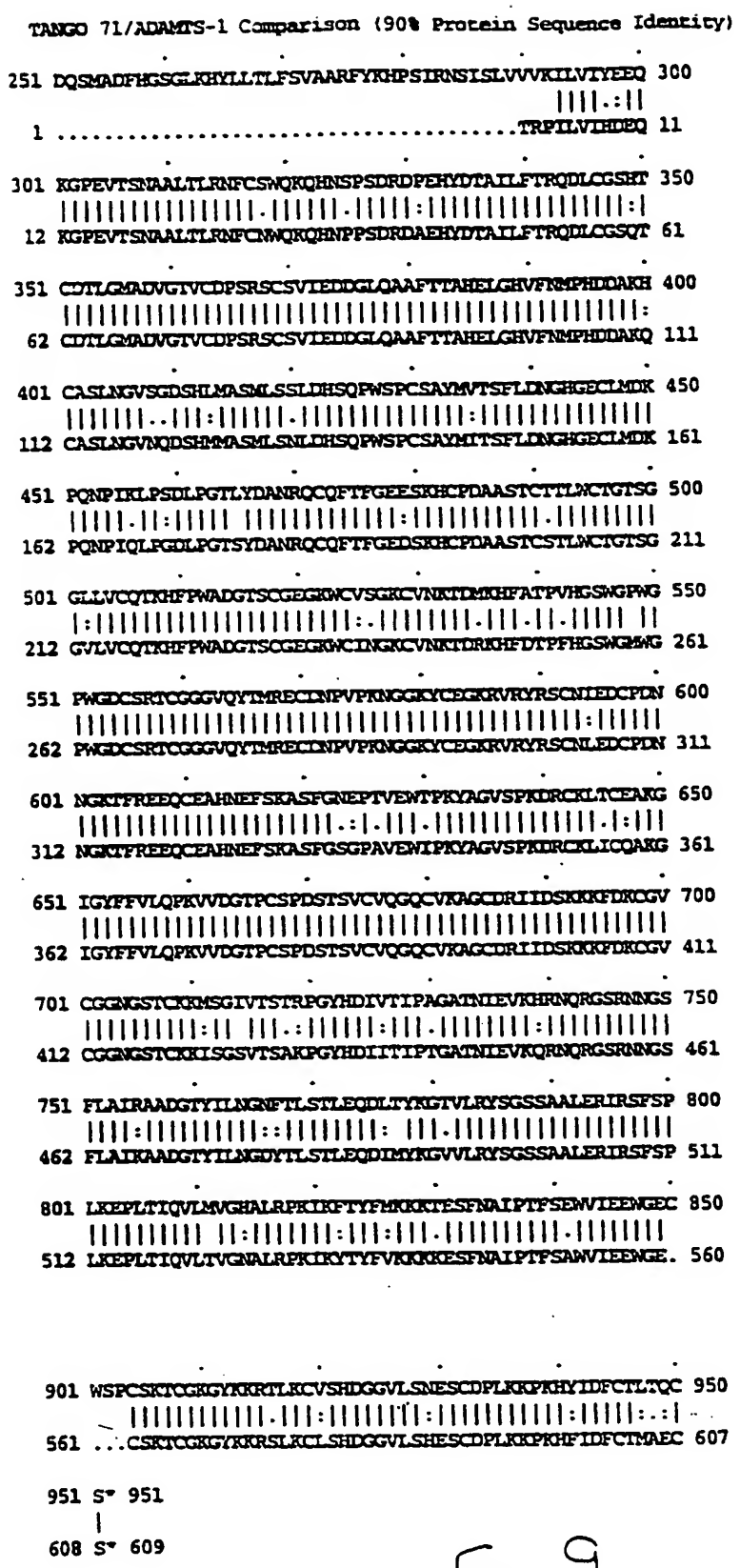

TANGO 71/ADAMTS-1 Comparison (90% Protein Sequence Identity)
251 DQSMADFHGSGLKHYLLTLFSVAARFYKHPSIRNSISLVVVKILVIYEEQ 300
1 .....TRPILVIHDEQ 11
301 KGPEVTSNAALTLRNFCSWQKQHNSPSDRDPEHYDTAILFTRQDLCGSHT 350
12 KGPEVTSNAALTLRNFCNWQKQHNPPSDRDAEHYDTAILFTRQDLCGSQT 61
351 CDTLGMADVGTVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKH 400
62 CDTLGMADVGTVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQ 111
401 CASLNGVSGDSHLMASMLSSLDHSQPWSPCSAYMVTSFLDNGHGECLMDK 450
112 CASLNGVNQDSHMMASMLSNLDHSQPWSPCSAYMITSFLDNGHGECLMDK 161
451 PQNPIKLPSDLPGTLYDANRQCQFTFGEESKHCPDAASTCTTLWCTGTSG 500
162 PQNPIQLPGDLPGTSYDANRQCQFTFGEDSKHCPDAASTCSTLWCTGTSG 211
501 GLLVCQTKHFPWADGTSCGEGKWCVSGKCVNKTDMKHFATPVHGSWGPWG 550
212 GVLVCQTKHFPWADGTSCGEGKWCINGKCVNKTDRKHFDTPFHGSWGMWG 261
551 PWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYRSCNIEDCPDN 600
262 PWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEGKRVRYRSCNLEDCPDN 311
601 NGKTFREEQCEAHNEFSKASFGNEPTVEWTPKYAGVSPKDRCKLTCEARG 650
312 NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQARG 361
651 IGYFFVLQPKVVDGTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDRCGV 700
362 IGYFFVLQPKVVDGTPCSPDSTSVCVQGQCVKAGCDRIIDSKKKFDRCGV 411
701 CGGNGSTCKKMSGIVTSTRPGYHDIVTIPAGATNIEVKHRNQRGSRNNGS 750
412 CGGNGSTCKKISGSVTSAKPGYHDIITIPTGATNIEVKQRNQRGSRNNGS 461
751 FLAIRAADGTYILNGNFTLSTLEQDLTYKGTVLRYSGSSAALERIRSFSP 800
462 FLAIKAADGTYILNGDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSP 511
801 LKEPLTIQVLMVGHALRPKIKFTYFMKKKTESFNAIPTFSEWVIEEWGEC 850
512 LKEPLTIQVLTVGNALRPKIKYTYFVKKKKESFNAIPTFSAWVIEEWGE. 560
901 WSPCSKTCGRGYKKRTLKCVSHDGGVLSNESCDPLKKPKHYIDFCTLTQC 950
561 ...CSKTCGRGYKKRSLKCLSHDGGVLSNESCDPLKKPKHFIDFCTMAEC 607
951 S* 951
608 S* 609

Fig. 9

COMPARISON OF TANGO 73 AND TMVCF
(SIROTKIN ET AL, GENOMICS 1997, 42:245-251)

PERCENT SIMILARITY: 67.619 PERCENT IDENTITY: 46.190

TMVCF.PEP X T73PRO JULY 29, 1997 16:32 ..

```
  1 MGSAALEILGLVLCLVGWGGLILACGLPMWQVTAFLDHNIVTAQTTWKGL 50
    |:.|:|::||::|.::|| | |:..:|| |.: .: :.||||||. :.||
  1 MANAGLQLLGFILAFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEGL 50

 51 WMSCVVQSTGHMQCKVYDSVLALSTEVQAARALTVSAVLLAFVALFVTLA 100
    ||||| ||||::||||:||:|.||..:||.||| | ::||:.:|:||. .
 51 WMSCVSQSTGQIQCKVFDSLLNLSSTLQATRALMVVGILLGVIAIFVATV 100

101 GAQCTTCVAPG.PAKARVALTGGVLYLFCGLLALVPLCWFANIVVREFYD 149
    | .| .|:..: ..| |:|:.||.::|:.||  ||: .|::| :|.||||
101 GMKCMKCLEDDEVQKMRMAVIGGAIFLLAGLAILVATAWYGNRIVQEFYD 150

150 PSVPVSQKYELGAALYIGWAATALLMVGGCLLCCGAWVCTGRPDLSFPVK 199
    | .||..:||:|.||:.||||..| ::||.||||    :..|.. |:|..
151 PMTPVNARYEFGQALFTGWAAASLCLLGGALLCC....SCPRKTTSYPTP 196

200 YSAPRRPTATGDYDKKNYV 218
     . |:.:...|    |:||
197 RPYPKPAPSSG....KDYV 211
```

Fig. 10

Comparison of TANGO 73 and RVP1

Percent Similarity: 68.750 Percent Identity: 48.077

t73pro x RVP1.pep July 29, 1997 17:38

```
  1 MANAGLQLLGFILAFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEGL 50
    . :|:: |  ||.|||:..||:.||| ||: .: |..|:|||  :|||
  1 .MSMSLEITGTSLAVLGWLCTIVCCALPMWRVSAFIGSSIITAQITWEGL 49

 51 WMSCVSQSTGQIQCKVFDSLLNLSSTLQATRALMVVGILLGVIAIFVATV 100
    ||.|| |||||:|||::||||.|. .|||.|||:||:|||:.::::|| |
 50 WMNCV.QSTGQMQCKMYDSLLALPQDLQAARALIVVSILLAAFGLLVALV 98

101 GMKCMKCLEDDEVQKMRMAVIGGAIFLLAGLAILVATAWYGNRIVQEFYD 150
    | .| .|: :||..| ::.:::|.:||||:: .||:..| :| |:.:||:
 99 GAQCTNCV.QDETAKAKITIVAGVLFLLAAVLTLVPVSWSANTIIRDFYN 147

151 PMTPVNARYEFGQALFTGWAAASLCLLGGALLCCSCPRKTTSYPTPRPYP 200
    |:.| ..: |:| :|:.|||||.| ||||||||||||.:.. ||. |.
148 PLVPEAQKREMGTGLYVGWAAAALQLLGGALLCCSCPPREKYAPTKILYS 197

201 KP...APSSGKDYV..................................... 211
     |   :|:.|.: .
198 APRSTGPGTGTGTAYDRKTTSERPGARTPHHHHYQPSMYPTRPACSLASE 247
```

Fig. 11

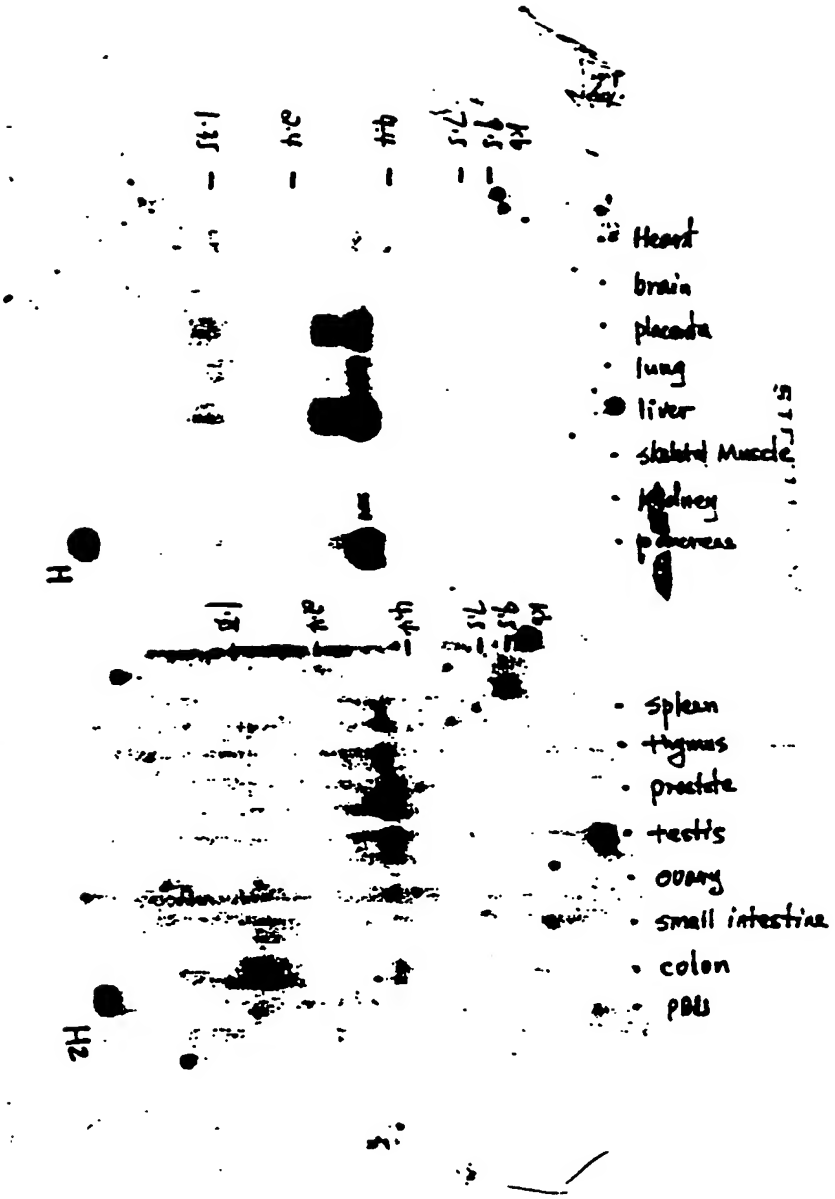
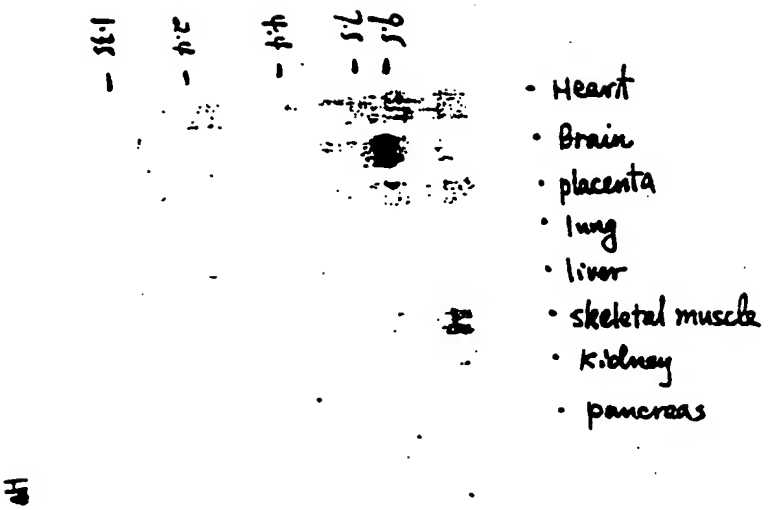

Heart
brain
placenta
lung
liver
skeletal Muscle
kidney
pancreas
spleen
thymus
prostate
testis
ovary
small intestine
colon
PBL

FIG. 12


9.5
7.5
4.4
2.4
1.35
Heart
Brain
placenta
lung
liver
skeletal muscle
kidney
pancreas

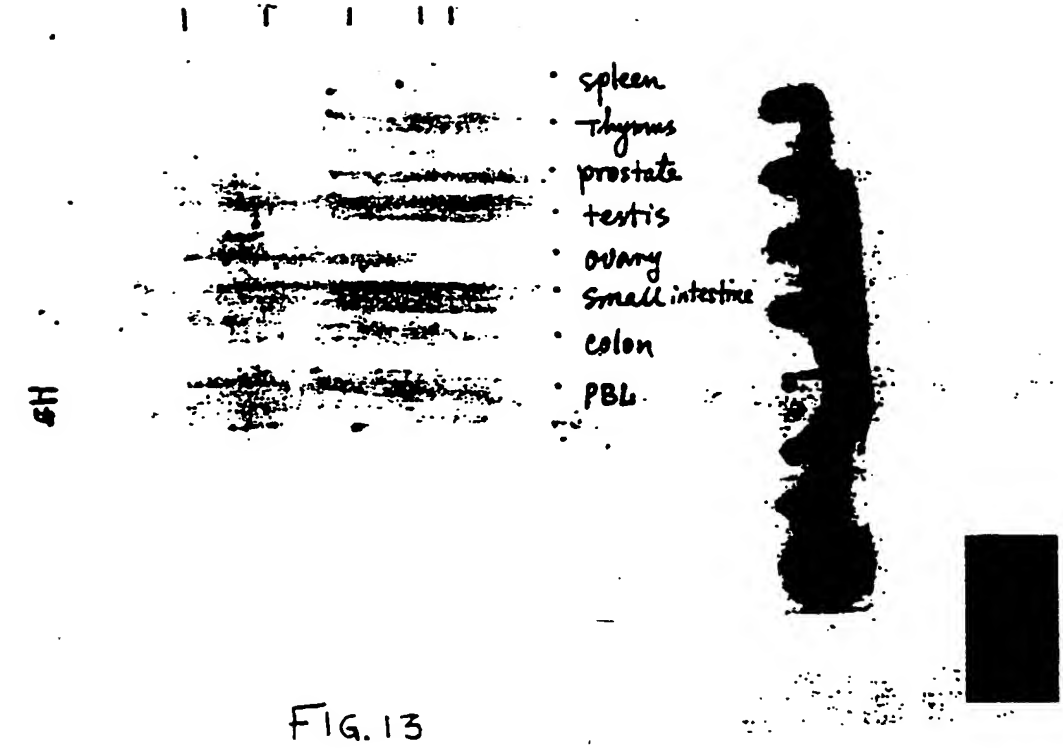

spleen
Thymus
prostate
testis
ovary
small intestine
colon
PBL

FIG. 13

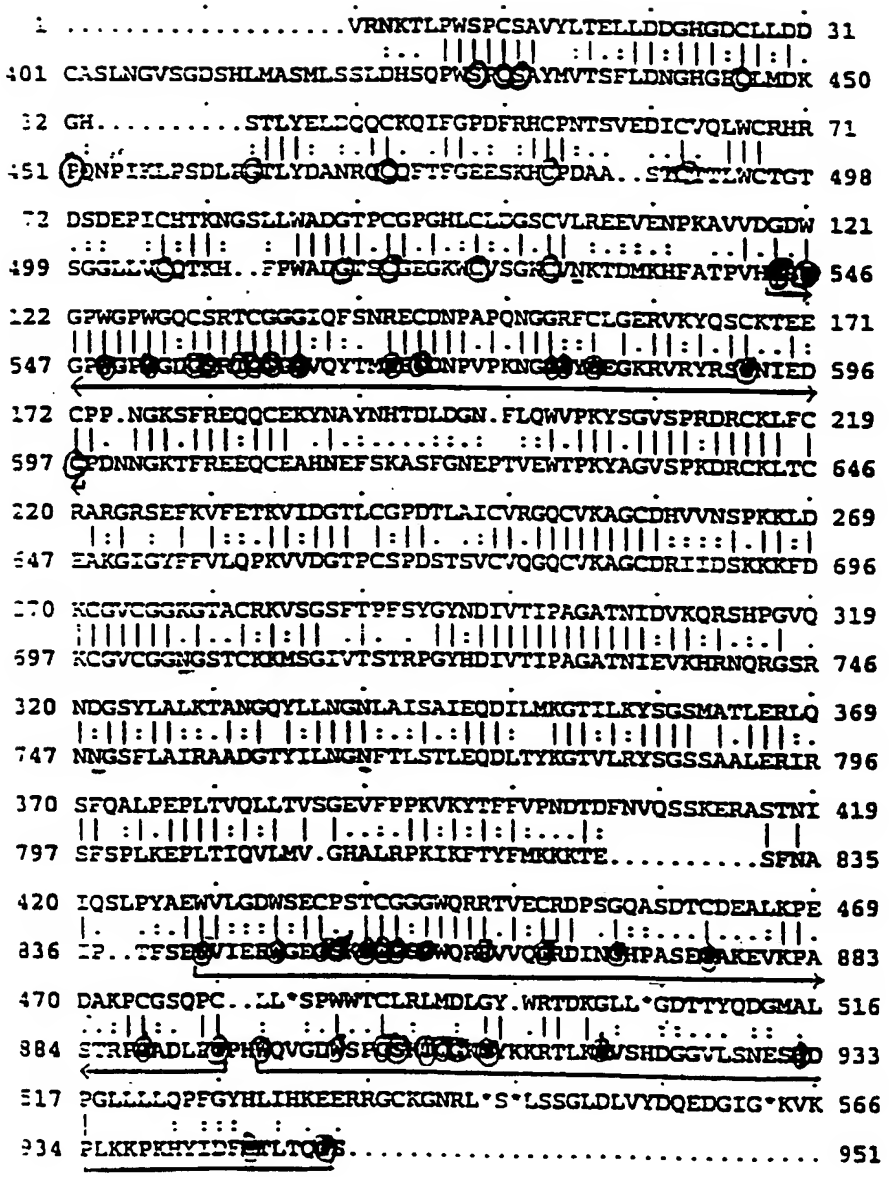


FIG. 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/16502

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 15/12, 15/16, 15/18, 15/63; C07K 14/46, 14/47, 14/475

US CL :530/350; 435/69.1, 71.1, 71.2, 325, 471, 252.3, 320.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 435/69.1, 71.1, 71.2, 325, 471, 252.3, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96/01896 A1 (HUMAN GENOME SCIENCES, INC.) 25 January 1996, see entire document.	1-10, 12
A	BRADHAM et al. Connective Tissue Growth Factor: a Cysteine-rich Mitogen Secreted by Human Vascular Endothelial Cells Is Related to the SRC-induced Immediate Early Gene Product CEF-10. The Journal of Cell Biology. September 1991, Vol. 114, No. 6, pages 1285-1294, see entire document.	1-10, 12

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 14 OCTOBER 1998	Date of mailing of the international search report 12 NOV 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer PREMA MERTZ Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/16502

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MASON et al. Dorsal Midline Fate In Drosophila Embryos Requires Twisted Gastrulation, A gene Encoding A Secreted Protein Related To Human Connective Tissue Growth Factor. Gene & Development. 1994, Vol. 8, pages 1489-1501, see entire document.	1-10, 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/16502

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, CAS ONLINE, MEDLINE, CAPLUS, EMBASE, BIOSIS
search terms: Tango-71, Tango-73, Tango-76, Tango-74, Tango-83, nucleic acid, DNA, polypeptide, protein, recombinant, cloning, production

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups invention which are so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must to paid.

Group I, Claims 1-10 and 12, drawn to Tango-71, Tango-73, Tango-74, Tango-76 and Tango-83 polypeptides, nucleic acid molecules encoding Tango-71, Tango-73, Tango-74, Tango-76 and Tango-83 polypeptides, a vector, a host cell, and a method for producing Tango-71, Tango-73, Tango-74, Tango-76 and Tango-83 polypeptides.

Group II, Claim 11, drawn to antibodies which bind Tango-71, Tango-73, Tango-74, Tango-76, and Tango-83 polypeptides.

Group III, Claims 13-15, drawn to a method for detecting Tango-71, Tango-73, Tango-74, Tango-76 and Tango-83 polypeptides in a sample.

Group IV, Claims 16-18, drawn to a method for detecting nucleic acid molecules encoding Tango-71, Tango-73, Tango-74, &Tango-76 and Tango-83 polypeptides in a sample.

Group V, Claims 19-22, drawn to a method for identifying a compound which binds to Tango-71, Tango-73, Tango-74, Tango-76 and Tango-83 polypeptides.

The inventions listed as Groups I-V do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Pursuant to 37 C.F.R.§ 1.475(d), the ISA/US considers that where multiple products and processes are claimed, the main invention shall consist of the first invention of the category first mentioned in the claims and the first recited invention of each of the other categories related thereto. Accordingly, the main invention (Group I) comprises the first-recited product, Tango-71, Tango-73, Tango-74, Tango-76 and Tango-83 polypeptides, nucleic acid molecules encoding Tango-71, Tango-73, Tango-74, Tango-76 and Tango-83 polypeptides, a vector, a host cell, and a method for producing Tango-71, Tango-73, Tango-74, Tango-76 and Tango-83 polypeptides. Further pursuant to 37 C.F.R.§ 1.475(d), the ISA/US considers that any feature which the subsequently recited products and methods share with the main invention does not constitute a special technical feature within the meaning of PCT Rule 13.2 and that each of such products and methods accordingly defines a separate invention.